Mathematical Models in Cancer

Systems Biology

Giuseppe Jordão

U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Department of Mathematics
University of Porto
Portugal, 2017

Mathematical Models in Cancer

Systems Biology

Giuseppe Jordão

Thesis presented for the degree of Master in Mathematical Engineering

Supervisor: Prof. Doutor João Nuno Tavares

Department of Mathematics
University of Porto
Portugal, 2017

To my mother and sister

Abstract

Nesta tese, são obtidos modelos matemáticos determinísticos a partir de modelos bioquímicos da célula eucariótica, em dois casos distintos, para efeitos de comparação: célula saudável e célula cancerígena. O primeiro modelo é baseado nos modelos de [3] e [8] e faz uso da cascata de sinalização MAPK, proposta em [7], e da via de transdução de PI3K/AKT descrita em [37], de modo a criar um modelo actualizado mais abrangente da regulação da célula saudável. O segundo modelo, da célula cancerígena, é construído a partir do primeiro modelo da célula saudável por alteração de vias específicas de transdução, e interpretando o resultado à luz da literatura em oncologia molecular. Esta interpretação é feita em duas abordagens: simulação de desregulações comuns e simulação de dois cancros específicos, o cancro do cólon e o cancro da mama.
O modelo actualizado proposto é analisado de acordo com a sua adequação aos modelos bioquímicos do ciclo celular, e ainda da sua robustez quando é usado para simular desregulações em cancro. São propostas terapias-alvo, de acordo com os resultados.
São, portanto, exploradas a eficácia e a utilidade da modelação matemática em fornecer resultados in silico, dos quais se possam retirar sugestões úteis para possíveis terapias. O objectivo é validar a matemática, novamente, como uma ferramenta poderosa com a qual se pode modelar a natureza dos sistemas biológicos e extrair conclusões para problemas da vida real.

Abstract

In this thesis, deterministic mathematical models are derived from biochemical models, within a human cell, in two distinct cases, for comparison: healthy cell and cancerous cell. The former model is based in [3] and [8] and makes use of the MAPK cascade pathway [7] and the PI3K/AKT pathway for signalling transduction [37], to create a wider updated model for the regulation of a healthy cell. The latter model, for the cancer cell, is derived from the healthy cell model, by altering specific pathways, and interpreting the outcome in the light of literature in cancer. This last study is done in two approaches: simulation of common deregulations and specific cancer simulation, colon and breast cancer. After studying both models, we propose targeting therapies and simulate their consequences. We thus explore mathematical modeling efficacy and usefulness in providing enough information from which to derive ideas for therapies. The purpose is to validate mathematics, once again, as a powerful tool with which one can model the underlying nature of biological systems and extract useful conclusions to real-life problems.

Contents

List of Figures

Introduction

Cancer is one of the most deadly diseases among humanity, in great part due to the large amount of variables which have to be taken into account in its development and dynamics, making it particularly difficult to approach therapeutically. Each cancer is unique in the sense that the particle's path arrangements, in the circumstances that originated it, can vary greatly according to the organism where it develops, the quality and quantity of its nurturing habits which, in turn, also depend on the resources of the environment where that organism lives, among many other degrees of freedom.
The understanding of how cancer mechanism works, starts with understanding how a healthy cell behaves, since the differences between cancer dynamics and healthy tissue dynamics are a reasonable object of analysis in cancer theory.
When a single mammalian cell fails to stop cell cycle, when it needs to, proceeding to replicate and originate offspring with anomalies, it can quickly develop a tumor whose priority is to grow and divide uncontrollably, selfishly, wearing all resources in its environment, destabilizing its neighbouring healthy cells in the tissue and, consequently, the whole organism. The study of individual healthy and cancerous cells dynamics is therefore an understandable approach for cancer therapy development and is the one we discuss in this thesis.
The advance in the technology relevant to this field, is itself divided in two main branches: improvement in computing power and in measurement accuracy. As the computing power increases, according to Moore's law, the ability to store huge amounts of data is enhanced and this is a feature of vital importance when it comes to simulate in a computer complex systems such as the ones embedded in cell metabolism or cancer morphogenesis. Of course all the storage capacity wouldn't mean much if there weren't breakthroughs in measurement tools. To accurately arm the simulations with reasonable predictions, the input information should be as close to reality as it can get. This is why the link between *in vitro*/*in vivo* research is of such importance for *in silico* approaches. In this way cancer, among other diseases, is being fought in many inter-connected battle fronts. The interaction between *in silico* and *in vitro* or *in vivo* technique, permits the upgrade of each one of these approaches without unnecessary waste of the resources. For instance, if the simulation of a specific cancer pathway exhibits a given feature for which it is known a treatment, *in vitro* or *in vivo* researchers can then focus on that specific feature, thus narrowing down the set of possible experiments, saving time and money in this process.
With this in mind, we will explore in this thesis an updated model for the healthy cell,

in order to simulate the cell cycle, and afterwords use it to study important pathways in cancer dynamics, whether they are commonly deregulated, relevant for cancer treatment, or because of the existence of cross-talks between them and cancer-related pathways. To accomplish these simulations, we will base our ground-knowledge on the general accepted chemical interactions within the cell, among biologists and biochemists, although this doesn't necessarily mean we completely exclude hypothesis which we find useful for the construction of the models.

The mathematics behind the creation of the models is reviewed in the first chapter, thus over viewing the contents of biomathematics and systems biology. The following chapters focus on the eukaryote cell cycle and the restriction point, p53 protein and signalling transduction pathways, i.e., in the biological and biochemical background needed to justify the reaction network chosen for modelling. We reserve chapter 5 to explore cancer's dynamics and main features and chapter 6 to summarize the state-of-the-art of mathematical models. Chapter 7 is dedicated to the explanation of our model of the cell cycle, which we then use to simulate the healthy cell in chapter 8 and common deregulations in chapter 9. In the last three chapters, two specific cancers are studied, colon cancer and breast cancer, as well as possible target therapies we could derive from the results of the simulations restricted to today's advances in personalized medicine.

Chapter 1

Preliminaries: Systems Biology

I Introduction and overview

In the 21st century, biology is being transformed from a purely lab-based science to a collaboration between *in vivo* or *in vitro* experiments and information science. As such, biologists have had to draw assistance from mathematicians, computer scientists and engineers. The result has been the development of the fields of **Bioinformatics** and **Computational Biology** (terms often used interchangeably).
The major goal of these fields is to extract new biological insight from the large noisy set of data being generated by high-throughput technologies. Initially, the main problems in bioinformatics were how to create and maintain databases for massive amounts of DNA sequence data. Addressing these challenges also involved the development of efficient interfaces for researches to access, submit and revise data. Bioinformatics has expanded into the development of software that also analyzes and interprets these data.
Systems Biology involves the collection of the large experimental data sets with which the development of mathematical models that predict important elements in this data is done. The quality assessment of these models by comparing numerical simulations with experimental data allows the update of these models to better fit observation. The ultimate goal of systems biology is therefore to develop models and analytic methods that provide reasonable predictions of experimental results. While it will never replace experimental methods, the application of computational approaches to gain understanding of various biological processes held the promise of helping experimentalists work more efficiently. These methods also may help gain insight into biological mechanisms when information could not be obtained from any known experimental methods. Eventually, it may be possible that such models and analytical techniques could have substantial impact on our society such as aiding in drug discovery.
System biologists analyze several types of molecular systems, including **genetic regulatory networks**, **metabolic networks** and **protein networks**. During the genomic age, standards for representing sequence data were (and still are) essential. Data collected from a variety of sources could not be easily used by multiple researchers without a standard

data format. For systems biology, standard data formats are also being developed. One format that seems to be getting some attraction is the *Systems Biology Markup Language* (**SBML**) - a **XML** based language for representing chemical reaction networks. This kind of modelling is implemented in software like **COPASI**.

The basic structure of a SBML model is a network consisting of a list of chemical species coupled with a list of chemical reactions. Each chemical reaction includes a list of *reactants*, *products* and *modifiers*, and also a mathematical description of the kinetic rate law governing the dynamics of this reaction.

For the purposes of this thesis, the reaction network consists of the three key elements: system components or pools of components, arrows that indicate flow of material and arrows that indicate flow of information or signals. By connecting pools with heavy arrows, we indicate which system components can be transformed into others, and following these arrows within the network, we obtain an impression of the different routes through which material can be processed by the system. In contrast to heavy arrows, we use dashed arrows to indicate that a system's component can affect or modulate a process in the system. An arrow of this type may represent, for instance, a *feedback inhibition* or the *activation* of a reaction.

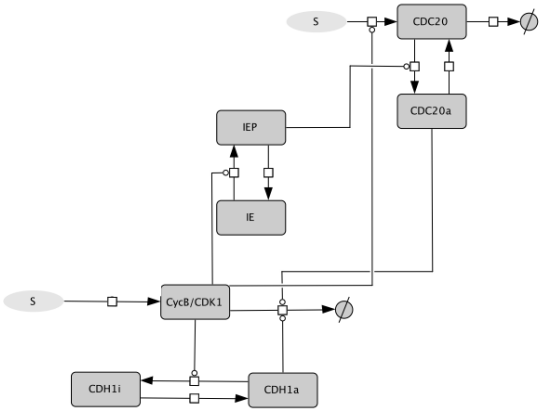

Figure 1.1: *Reaction network. SBML representation in Cell Designer.*

The overall strategy of analyzing a biochemical system consists of a sequence of six steps:

- List all components or pools of components that affect the system.
- List all interactions between these components and all modulations by which components affect the system.
- Arrange components, pools, interactions and modulations in the form of a network.
- Transcribe the network in terms of mathematical symbols and equations.
- Analyze these equations.
- Interpret the results.

The list of steps signals the fundamental importance of setting up the network in a proper form. The network connects reality and mathematical analysis, and if this connection is faulty, the results are unreliable or even wrong. When all components and interactions of a biochemical system are known, and when we strictly adhere to a few rules, it is usually not too difficult to construct a proper network.

On the other hand, ill-defined components, using the wrong types of arrows or confusing the flow material with a regulatory influence often leads to incorrect conclusions. In some cases, one may detect those problems early on, however, once the equations are formulated - correctly of incorrectly - it is often very difficult to detect inconsistencies in the map or the equations until it is time to interpret the results.

II The cell as a system

The functioning of a cell requires multiple processes to work in an orchestrated manner. Basic properties of cellular life, such as proliferation, macromolecule synthesis and degradation, and cellular metabolism, have to be tightly controlled. Failure in the regulation of these cellular functions, for example through mutations of specific genes, can result in another cellular phenotype and, eventually at the organism level, in severe diseases such as cancer. Whatever the cell type - eukaryotic or prokaryotic - or its neighboring environment - other cells, for example, like in a multicellular organism, or other neighborhood, like in unicellular organisms - its ability to respond to external stimuli, derive energy and materials needed to continually fabricate itself and eventually reproduce, is always controlled by a complex network of chemical reactions. Each reaction needs to be catalyzed by specific proteins, which are specialized molecules produced by the cell itself. Many proteins are enzymes, which are biological catalysts present in nearly every activity of the cell. Other proteins are used as structural elements to build cellular parts, as activation or repression agents to control reactions, as sensors to environmental condition, or take part in one of the many other tasks necessary for cellular function. There are thousands of different proteins in each cell. Producing these proteins not only requires the cell to obtain energy and materials, but also requires detailed communication between different parts of a cell or between cells. Much of the cellular machinery is devoted to ensuring the production of proteins at the right moment, in the right quantity, in the right place.
A cell's most reliable way to pass on the recipe for making proteins is contained in its genetic material and is passed on to daughter cells at each division. The machinery for reading this information is one of the core components of all living things and is highly similar in all types of cells, being constituted by a complex of enzymes. The information itself, called genetic material, is formed by molecules of DNA (deoxyrribonucleic acid), which have a sequential structure that enables them to act as information storage devices.

III Chemical kinetics

As mentioned in the last section, the metabolism of a single cell can be broken down to a network of chemical reactions. In this section we will show how to derive a system of differential equations from a set of chemical reactions and then proceed to analyze the resulting model quantitatively following the steps listed in section I.
A **Reaction Network** is a set of reactions $R_m u,\ \mu = 1, \cdots, r$, between reagents and products from the set of involved species, $\{S_i\}_{i=1,\ldots,s}$.

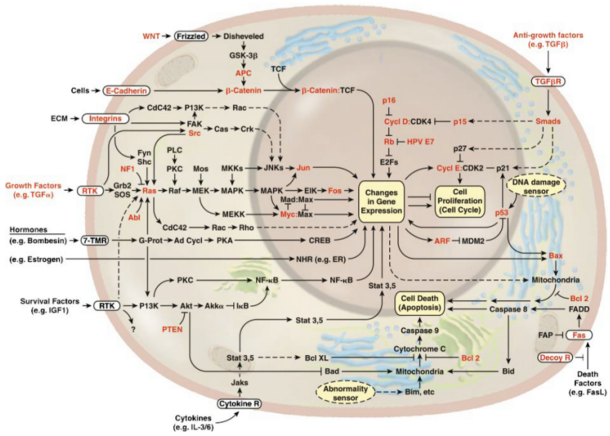

Figure 1.2: *The cell mechanisms. From Douglas Hanahan and Robert A. Weinberg The Hallmarks of Cancer, Cell. Vol. 100, 57-70, 2000*

$$R_\mu : \sum_{i=1}^{s} \alpha_{i\mu} S_i \xrightarrow{k_\mu} \sum_{i=1}^{s} \beta_{i\mu} S_i, \qquad \mu = 1, 2, ..., \tag{1.1}$$

with s species and r reactions. The scalars $\alpha_{i\mu}$ and $\beta_{i\mu}$ are called the stoichiometric coefficients, representing the participation of each species S_i, respectively as reagent and as product in reaction R_μ. The rate constant k_μ gives information on the kinetics of this reaction. The **Stoichiometric Matrix** is defined by the reaction network as:

$$S_{i\mu} = \beta_{i\mu} - \alpha_{i\mu}, \quad i = 1, ..., s; \mu = 1, ..., r \tag{1.2}$$

To study the dynamics of the reaction network (1.1), we define the variables:

- $N_i(t)$ = number of molecules of species S_i (as a reagent or as a product) present in instant time t.
- $Z_\mu(t)$ = number of occurrences of reaction R_μ in time interval [0, t].

In the time interval [0, t], each reaction R_μ occurs $Z_\mu(t)$ times. Each occurrence adds the amount $\mathbb{S} = \beta_{i\mu} - \alpha_{i\mu}$ to the current number of molecules of species S_i. Therefore, the number of molecules of species S_i at time t can be writen as

$$N_i(t) = N_i(0) + \sum_{\mu=1}^{r} \mathbb{S}_{i\mu} Z_\mu(t), \quad i = 1, ..., s \tag{1.3}$$

The term $N_i(0)$ is the number of molecules of species S_i at time $t = 0$, and it is a constant value for each i since the system of reactions considered is assumed to be closed, i.e., there are no additions nor losses of any species at any time. This value may increase or decrease throughout time according to the sum $\sum_{\mu=1}^{r} \mathbb{S}_{i\mu} Z_\mu(t)$.

Using vector notation, $\mathbf{N}(t) = (N_1(t), ..., N_s(t))^\top \in \mathbb{Z}_+^s$, $\mathbf{X}(t) = (X_1(t), ..., X_s(t))^\top \in \mathbb{R}_+^s$, $\mathbf{Z}(t) = (Z_1(t), ..., Z_s(t))^\top \in \mathbb{Z}_+^r$, we can write (1.3) as

$$\mathbf{N}(t) = \mathbf{N}(0) + \mathbb{S}\mathbf{Z}(t) \tag{1.4}$$

Dividing (1.4) by the mass volume, Ω, we obtain the equation in terms of concentrations

$$\mathbf{X}(t) = \mathbf{X}(0) + \frac{\mathbb{S}\mathbf{Z}(t)}{\Omega} \tag{1.5}$$

However, reactions are events that occur in a discrete set of time instants. Moreover, the time of occurrence of a reaction and the reactions that occur in the set of possible reactions are random variables since they are determined by several microscopic factors. Therefore a deterministic description has to be based on several simplifying assumptions:

- Reactions are so frequent that the number of occurrences $Z(t)$ can be approximated by a continuum variable $z(t)$. This assumption requires that a large number of molecules are in a large volume free to interact with each other. It also requires that the physical characteristics of each molecule (energy, orientation, etc.) favor the interactions which translate into a rate constant k. The presence of a large number of molecules also means that the occurrence of a reaction translates into a small change of N_i, so we can also approach $N(t)$ by a continuous variable $n(t)$. The concentration $X(t)$ can also be approximated by a continuous $x(t) = X(t)/\Omega$.

With these assumptions, (1.4) and (1.5) lead respectively to

$$\mathbf{n}(t) = \mathbf{n}(0) + \mathbb{S}\mathbf{z}(t) \tag{1.6}$$

$$\mathbf{x}(t) = \mathbf{x}(0) + \frac{\mathbb{S}\mathbf{z}(t)}{\Omega} \tag{1.7}$$

Taking the derivative over time, we get

$$\dot{\mathbf{n}}(t) = \mathbb{S}\dot{\mathbf{z}}(t) \tag{1.8}$$

$$\dot{\mathbf{x}}(t) = \frac{\mathbb{S}\dot{\mathbf{z}}(t)}{\Omega} \tag{1.9}$$

However, these ordinary differential equations are only useful if we can establish a relationship between the derivative $\dot{z}(t)$ and the variables n or x. Suppose that it is possible to establish such a relationship, $\dot{z}(t) = \hat{v}(n) = \Omega v(x)$. Thus, (1.8) and (1.9) could be written as $\dot{n}(t) = \mathbb{S}\hat{v}(n)$ and $\dot{x}(t) = \mathbb{S}v(x(t))$.
In many reactions, the rates v and $\hat{v}$ are proportional to the products of powers of concentrations of the reagents, elevated to a certain exponent (mass-action law, Gulberg and Waage 1864-1879):

$$\hat{v}_\mu = \hat{k}_\mu \prod_{i=1}^{s} n_i^{\alpha_{i\mu}}; \;\; v_\mu = k_\mu \prod_{i=1}^{s} x_i^{\alpha_{i\mu}} \tag{1.10}$$

Substituting (1.10) in $\dot{z}(t) = \hat{v}(n) = \Omega v(x)$, we get

$$\hat{k}_\mu \prod_{i=1}^{s} n_i^{\alpha_i\mu} = \hat{v}_\mu(n) = \Omega v_\mu(x) = \Omega k_\mu \prod_{i=1}^{s} x_i^{\alpha_i\mu}$$

As $x = n/\Omega$, then

$$\hat{k}_\mu \prod_{i=1}^{s} n_i^{\alpha_{i\mu}} = \frac{\Omega}{\Omega^{\sum_{i=1}^{s} \alpha_{i\mu}}} k_\mu \prod_{i=1}^{s} n_i^{\alpha_{i\mu}} \leftrightarrow \hat{k}_\mu = \frac{k_\mu}{\Omega^{\sum_{i=1}^{s} \alpha_{i\mu} - 1}} \tag{1.11}$$

IV Enzymatic Michaelis-Menten kinetics

Enzyme kinetics studies the chemical reactions catalyzed by enzymes, particularly the rate of reaction. The study of the kinetics of an enzyme allows elucidate the details of its catalyst mechanism, its role in metabolism, how its activity is controlled in the cell, for example how it can be inhibited by drugs or poisons or potentiated by other molecules. Many chemical and biological systems rely on enzymes that catalyze (i.e., accelerate the rate of a reaction without being consumed during the process), one or more of the possible reactions. A relatively simple model was proposed by Michaelis and Menten [45], where the reaction involves an enzyme E manipulating a substrate S which in turn reacts to form a new molecule, the product P [figure 1.3].

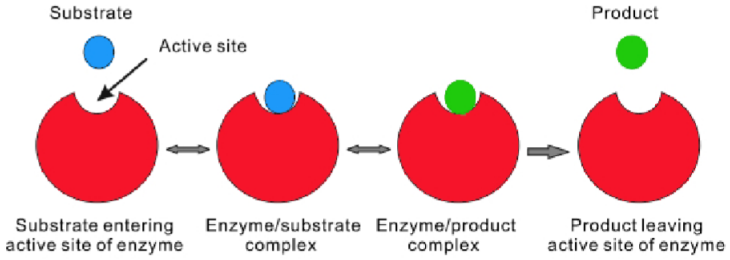

Figure 1.3: *Michaelis-Menten model*

The reaction network is

$$\underbrace{S + E \underset{k_d}{\overset{k_b}{\rightleftharpoons}} SE}_{binding/dissociation} \underbrace{\xrightarrow{k_{cat}} P + E}_{catalysis} \tag{1.12}$$

Using mass-action law and the notation for concentrations $x_E = E$, $x_{SE} = SE$, $x_P = P$, the kinetics equations are

$$\begin{cases} \frac{dS}{dt} = -k_b S \cdot E + k_d SE \\ \frac{dSE}{dt} = k_b S \cdot E - (k_d + k_{cat})SE \\ \frac{dE}{dt} = -k_b S \cdot E + (k_d + k_{cat})SE \\ \frac{dP}{dt} = k_{cat} SE \end{cases} \tag{1.13}$$

As initial conditions, assume there is a certain quantity of S and E, but no SE nor P, i.e., $S(0) = S_0$, $E(0) = E_0$, $SE(0) = 0 = P(0)$. Two useful conservation laws are $\frac{d}{dt}(E + SE) = 0$ and $\frac{d}{dt}(S + SE + P) = 0$ which leads to $E + SE \equiv E_{total} = E_0$, $S + SE + P \equiv S_0$. The first equation tells us that enzyme E is not produced or consumed during the reaction, it can be free or part of complex SE, but its total concentration remains constant equal to E_{total}. Therefore we can reduce the four kinetic equations above, only to two

$$\begin{cases} \frac{dS}{dt} = -k_b E_{total} \cdot S + (k_d + k_b)SE \\ \frac{dSE}{dt} = k_b E_{total} \cdot S - (k_{cat} + k_d + k_b S) \cdot SE \end{cases} \tag{1.14}$$

However, these equations are intractable. One of the most used simplifying hypothesis (Briggs and Haldane) is to assume that, with the progress of the reaction dynamics, the system reaches a quase-steady state in which the concentration of the complex SE remains constant. Experimental studies have shown that the concentration of the complex SE reaches a steady state much faster than the substrate. At equilibrium, the concentration SE doesn't change with time. Therefore, we can assume that $\frac{dSE}{dt} = 0$. From the second equation in (1.13), we get

$$SE = \frac{E_{total}S}{K_m + S} \tag{1.15}$$

where $K_m = \frac{k_d + k_{cat}}{k_b}$ is the Michaelis-Menten constant. For the reaction rate we obtain the Michaelis-Menten equation

$$v = \frac{dP}{dt} = k_{cat}SE = \frac{k_{cat}E_{total}S}{S + \frac{k_d + k_{cat}}{k_b}} = \frac{V_{max}S}{K_m + S} \tag{1.16}$$

where $V_{max} = k_{cat}E_{total}$.

The reaction network established in the model, explained in chapter 6, follows mass-action law and Michaelis-Menten kinetics, as well as Hill's equation.

V Cooperation. Hill equation

Many enzymes have more than one site to bind to substrates. If these sites act independently of one another, the enzyme is said to be **non cooperative**. If the binding of a molecule of substrate to a site facilitates the attachment of another molecule at a second site, the enzyme is said to be **positively cooperative**. If, however, the binding of a substrate molecule at a site hinders the binding of another molecule at a second site, the enzyme is said **negatively cooperative**.

Suppose that an enzyme can bind to two molecules of S substrate and that both bindings allow for the synthesis of the same product P. The generic name for a binding molecule is **Ligand** (Latin: *Ligare*). The binding reaction can be represented by:

$$\begin{aligned} S + E &\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} SE \xrightarrow{k_2} P + E \\ SE + S &\underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} SSE \xrightarrow{k_4} P + SE \end{aligned} \tag{1.17}$$

Using the conservation law

$$E + SE + SSE = E_{\text{total}}$$

and the quasi-stationary

$$\frac{dSE}{dt} = 0 = \frac{dSSE}{dt}$$

we can deduce for the reaction rate $S \longrightarrow P$, $v_{S\to P} = \frac{dP}{dt} = \frac{(k_2K_2 + k_4S) \cdot E_{\text{total}} \cdot S}{K_1K_2 + K_2S + S^2}$, where $K_1 = (k_{-1} + k_2)/k_1$ and $K_2 = (k_{-3} + k_4)/k_3$. Let's examine two cases:

- First let's consider the case of non-cooperation – the binding sites act independently and identically. So $k_1 = 2k_3 = 2k_+, 2k_{-1} = k_{-3} = 2k_-$ e $2k_2 = k_4$ where k_+ e k_- are reaction rates for binding to each of the sites individually. Doing $K = \frac{k_- + k_2}{k_+}$, we obtain $K_1 = K/2$, $K_2 = 2K$ and so

$$v = \frac{2k_2 E_{\text{total}}(K + S) \cdot S}{K^2 + 2K + S^2} = \frac{2k_2 E_{\text{total}} \cdot S}{K + S}$$

  i.e. the reaction rate with two binding sites is exactly twice the speed for only one place.
- Next, consider the case of positive cooperation, where the first substrate binding molecule is slow and this first connection turns the second faster. Keeping $k_1 k_3$ constant, if $k_1 \to 0$, then $k_3 \to \infty$, . In this case, reaction speed is

$$v = \frac{V_{\max} S^2}{K_m^2 + 2K + S^2}$$

  where $K_1 K_2 = K_m^2$ and $V_{\max} = k_4 E_{\text{total}}$.
  In general, if the enzyme has n binding sites, then

$$v = \frac{V_{\max} S^n}{K_m^n + 2K + S^n}, \qquad \textbf{Hill equation} \tag{1.18}$$

  The exponent n is generally determined from experimental data and can be non-integer, therefore not equal to the number of active sites. For a simple proof of the generalization for n binding sites, see [55].

Chapter 2

The Cell Cycle: An overview

I Different phases

Cell proliferation involves the reproduction of a cell to originate two daughter cells, each with the potential to originate their own offspring. This activity in multicellular organisms is fundamental not only to produce cells essential for development and growth, but also to replace cells as they die, and it functions in a periodical fashion, following a sequence of stages. This periodic sequence is called the **cell cycle**.
The cell cycle of eukaryotic cells can be divided in two main events: replication of DNA, known as S phase, and mitosis, known as M Phase, followed by cytokinesis. Between S phase and M phase the cell enters G_1 and G_2 phase, in which different concentrations of biomolecules change. The set of phases which includes G_1, S and G_2 phases is called **interphase** . When the cell is not in mitosis nor in interphase, it means it is in a quiescence state, the so called G_0 phase, or is preparing itself for apoptosis, i.e. programmed cell death.
After cell division, each one of the daughter cells enter G_2 phase, in which they prepare to enter mitosis. Mitosis proceeds through a series of stage conversions characterized by the location and behaviour of the chromosomes. Some of the conversions during mitosis are irreversible transitions. The first stage during mitosis is called **prophase**, in which the chromosomes within the nucleus rearrange themselves to become condensate. In warm-blooded creatures with small chromosomes, this stage can last less than 15 minutes. At some point of prophase, the cell commits to mitosis, i.e., it passes through a series of biochemical changes which are irreversible transitions. Before this point in prophase is reached, chromosome condensation can be reversed by physical or chemical alterations to the cell. Prophase in eukariotic cells is also commonly marked by the appearance of two centrosomes, which is a molecular arrangement of microtubules whose function is to help stabilizing the structure of the cell. The decomposition of the nuclear envelope signals the beginning of the **prometaphase**. During prometaphase, the chromosomes interact with the two centrosomes and their associated arrays of microtubules to form the spindle. As the chromosomes become attached to the spindle, they go through a series of complex

motions called congression, which make up the movement of the chromosomes to a plane at the "spindle equator", halfway between the two poles. Even though complex, the events that make up congression are reversible. In most cells, prometaphase is the longest stage of mitosis, since it lasts until all of the chromosomes are positioned at the equator. This may take just a few minutes in embryos or up to several hours in highly flattened tissue cells.
Once the chromosomes are all lined up in the spindle equator, the cell is said to be in **metaphase**. Metaphase ends when the two sister chromatids of each chromosome separate, beginning **anaphase**. The separation of chromatids in the beginning of **anaphase** marks another point-of-no-return in mitosis, because after the separation the "glue" holding together the chromatids is destroyed. Following separation, each sister chromatid moves towards a different pole of the spindle. The two poles themselves move farther apart. As the two groups of chromosomes move apart, the spindle disassembles.
The final stage of mitosis, **telophase**, begins when each of the two groups of chromosomes start forming their own nucleus (see figure 2.1).

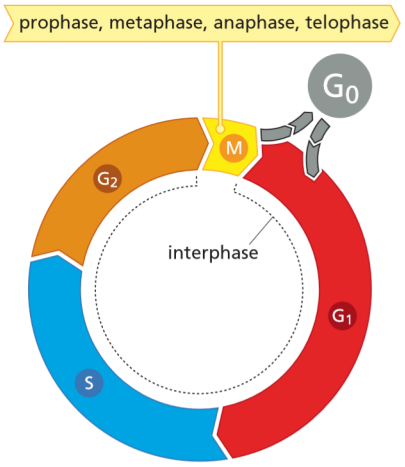

Figure 2.1: *The cell cycle (Weinberg)*

After mitosis is successfully over, the cell is ready to divide. Eukariotic cells perform this crucial event in reproduction, by constricting between the two newly separated sets of chromosomes in a phenomena called **cytokinesis**. The cell then enters G_1 phase, in which it prepares for DNA replication phase by synthesizing mRNA. In this phase, a very important point-of-no-return was identified, called the **restriction point**. The restriction point is explored in more detailed in chapter 3. The phase following G_1 phase is the S phase, during which the DNA is duplicated. The cell then enters G_2, growing and preparing itself for another round of mitosis, hence completing the cycle.
The cell cycle's average duration is 16 hours (15 hours for interphase and 1 hour for mitosis), varying according to cell type.

In an adult, more than 25 million cells undergo cell division per second. The magni-

tude of this number justifies the need for a precise mechanism of regulation of the cell cycle. The biomolecules that regulate this cycle are the **cyclins** (Cyc's) and the **cyclin dependent kinases** (CDK's), proteins and enzymes, respectively. The concentration of cyclin proteins is dependent on the transcription of its gene and by subsequent regulated protein degradation. The pairing of cyclins to the CDK's is highly specific. Cyclins are regulatory subunits of their CDK's, and upon binding of a cyclin to its CDK partner, the cyclin induces a conformational change in the catalytic subunit of the CDK revealing its active site. Different Cyclin/CDK complexes are present at specific phases in the cell cycle, and they are important regulators of irreversible phase transitions.
To enter the cycle from G_0 phase, some external signal must be transducted through the cell's cytosol reaching the nucleus and promoting transcription of CycD and CDK4,6, thus conducing the cell to enter G_1 phase. This external signal is transducted via signalling pathways, two of which are explored in chapter 5.
Cyclin D plays a role in the regulation of expression of the cyclin E gene, and consequently, during the transition between G_1 and S phase, CycE/CDK2 complexes increase their concentration in the cytosol, allowing for the transcription of CycA and CDK2, which, in the form of complex, promotes the movement to the G_2 phase of the cycle, where CycA/CDK1 complexes are predominant, leading the passage to mitosis, where, in turn, CycB/CDK1 complexes are in abundance. Activation of the **anaphase-promoting complex** (APC) by binding of **cell-division cycle protein 20** (CDC20) and **cadherin 1** (CDH1) is necessary for exiting mitosis. This completes the cycle of concentrations of Cyclin/CDK complexes, right before the cytokinesis event.
Two families of inhibitors are involved in regulating Cyclin/CDK activity: the $p16^{\text{ink4a}}$ (INK) family and the p21 (Cip/Kip) family. The INK proteins bind CDK's 4/6 and interfere with the binding of CDK's 4/6 to Cyclin D. The p21 proteins interact with both cyclins and their associated CDK's, blocking the ATP-binding site, thus disabling this site, blocking kinase activity (hence the designation *inhibitor*).

II Cell cycle arrest and Apoptosis

The cell cycle can be disrupted by the cell itself if something is not according to the regulation we summarized in the previous sections. As it was mentioned in the last section, the cell cycle is armed with specific checkpoints to prevent anomalies in development by correcting or stopping any malformations or errors before committing to following phases in the cycle. An extensively studied checkpoint is the G_1 checkpoint which, for the purpose of this thesis, will be explored in more detail in the following section. Another one of these checkpoints can be found during G_2 phase, after S-phase and before mitosis. When a G_2 cell was fused with an S-phase cell, the G_2 phase nucleus "waited" for the S-phase nucleus to finish DNA replication before undergoing nuclear envelope breakdown and entering mitosis. This suggested that a mechanism exists to prevent mitosis until DNA replication is complete. There is also a checkpoint that monitors DNA replication process during S-phase.
Even though these surveillance mechanisms happen during distinct phases of the cycle, they all consist of a sensor that detects a defect in an event, a signalling module that

transmits a signal upon detection of an error, and a target that is part of the cell cycle engine controlled to halt cell cycle progression. In some cases, this conduces the cell to a specific fate called programmed cell death. In general, it occurs by the mechanism of cell death denominated **apoptosis**. Although most programmed cell deaths occur by apoptosis, these are not the same thing. Programmed cell death refers specifically to cell death occurring at a defined point in development, whereas apoptosis is defined by morphological features of the cell death. Apoptosis is a mechanism of defense developed to protect multicellular organisms from malformations in cell development and/or activity, for it conduces the cell to destroy itself without damaging neighboring cells. It does so by shrinking, condensing, tearing up its outer layers and breaking the DNA into fragments [10]. However, apoptosis must be carefully regulated. Inadequate activation of apoptosis can originate the unnecessary destruction of cells seen in some neurodegenerative diseases. It is important to note that not all cell death is apoptotic. It can also happen due to necrosis,i.e., disabling damage or trauma, making it impossible for the cell to survive. Apoptosis, on the other hand, is stimulated when the DNA is damaged, the cell suffers withdrawal of essential growth factors or nutrients or it is attacked by cytotoxic lymphocyte. When DNA is damaged, the ATM ("ataxia-telangiectasia mutated") kinase is activated, culminating in p53 concentration increase, which in turn gives place to a sequence of events that turn on Caspase-9 and ultimately induces apoptosis. The details on this mechanism are far too extensive for the purpose of our model, which is why we kept the apoptotic dynamics fairly simple, as we explain in chapter 7.

III The Restriction Point Regulation

In the late G_1 phase there is some device that allows the cycle to continue regardless of mitogenic activity at the membrane. This point, called **Restriction Point**, was set between the 3rd and the 4th hour of G_1 phase [11], and its regulation is managed by the expression of the **retinoblastoma protein**, Rb.
Retinoblastoma is a rare childhood cancer. It is caused by a mutation in the retinoblastoma tumor supressor gene, named after the disease, and found in chromosome 13 [46]. The retinoblastoma protein, Rb, whose transcription is done from the retinoblastoma gene, plays an important role in regulating the restriction point along with E2F transcription factor family, which is crucial for the expression of genes needed for S phase. Active E2F, migrates to the nucleus of the cell where it promotes DNA replication, initiating S phase. Active Rb interferes with the transactivation domain of E2F, deactivating E2F and thus inhibiting the passage to S phase. Rb is activated in its hypophosphorylated form and is deactivated in its phosphorylated form. In response to a growth signal, CycD/CDK4,6 complexes inhibit active Rb, phosphorylating it partially, leading to a partial activation of E2F. PP1 phosphatase dephosphorylates Rb, increasing the concentration of active Rb and thus promoting the inhibition of E2F. Along with CycE/CDK2 complexes, E2F promotes the passage through the G_1-S phases frontier, hence leading to DNA replication, independent of further mitogenic signals. At this point, the cell enters in automatic program.

Chapter 3

P53: The Guardian of the Genome

I P53 pathway

The p53 gene, found in 1979 by separate groups of investigators [5-8], and set to be a **tumor supressor gene** in 1989 ([17], [18]), expresses the p53 protein, a central biomolecule in cancer research, specifically in the study of pathways within the cell. This is due to the fact that virtually all cancers exhibit some sort of mutation of p53 gene or modifications to its pathway. The study of p53 pathway revealed the core of its regulation as well as several links that it establishes between other major pathways, such as the one of Rb protein, E2F and Ras. The concentration of p53 protein within an unstressed cell is low, however it has a fast turnover when the cell is under stress or has suffered DNA damage. Upstream stress activators include radiation-, drug- or carcinogen-induced DNA damage. P53 can elicit downstream cellular effects, which include transient or permanent cell cycle arrest, DNA repair, apoptosis, and inhibition of angiogenesis. The ability it has to induce cell cycle arrest allows for the repair of DNA damage. Cell cycle inhibition and apoptosis are, however, two independent effects of p53.
The core regulation of p53 protein is co-protagonized by the protein Hdm2 (Mdm2 in the mouse) that inhibits p53 protein by binding to it directly. P53 protein promotes the transcription of Hdm2, defining a negative feedback loop between p53 protein and Hdm2 [19], [20]. $p14^{ARF}$ ($p19^{ARF}$ in the mouse) in turn inhibits Hdm2 and its activity is inhibited by p53 protein. The transcription factor E2F also plays a role in p53 regulation, by sustaining a negative feedback loop with $p14^{ARF}$ by inducing it while being inhibited by it [9]. In [21], downstream events were explored in distinct pathways, as well as useful positive and negative feedback loops for p53 protein.
Inhibition of cell cycle involves the transcriptional induction of p21 gene. Its product, the p21 protein, inhibits several Cyclin/CDK complexes and causes a pause in the G_1 to S (and G_2 to mitosis) transition of the cell cycle.
Let us resume the main downstream event triggered by p53 protein activity which culminates in cell cycle arrest: the p21 gene product, a Cyclin Dependent Kinase Inhibitor (CKI), that inhibits CycE/CDK2 complex is a relevant molecule in p53-mediated G_1-S phase

arrest. Its transcription is induced by p53 protein activity. In addition, p21 also binds PCNA (proliferating cell nuclear antigen), a protein that has a role in DNA synthesis and DNA repair. The interation with p21 inhibits PCNA's role in DNA replication. This facilitates the action of p53 in stopping cell cycle. The regulation of p21 gene is therefore important in p53 decision-making process.
There is also CDC25 inhibited by 14-3-3-sigma, and CDC2 induced by CDC25 and CycB, the latter inhibited by Gadd45. CDC2 promotes Cell Cycle arrest between G_2-S phase. This last pathway is not of our interest, as it concerns another checkpoint in the cell cycle, not the restriction point. The cell cycle arrest pathway in which we focused our attention was the one concerning the checkpoint during G_1-S phase transition, and is obviously of the most relevance for studying the regulation of the restriction point.
Cysteine aspartate-specific proteases (caspases) regulate many of the cellular and biochemical changes in the dying cell undergoing apoptosis. They do not need to be newly synthesized upon activation of apoptosis, as they are present in inactive forms. Executioner, initiator and inflammatory caspases are the three main types of caspases. Executioner caspases are responsible for cleaving many different proteins, and it has been estimated that there are approximately 500 substrates for caspases in mammalian cells. In most cases of apoptosis, the cleavage of these executioner caspases is mediated by another set of caspases, the initiator caspases. The activation of the executioner caspases by the initiator caspases defines the different pathways of apoptosis.
Another important element in the initiation of apoptosis is cytochrome c, which is one of the components of the electron transport chain in the mitochondria. It has the ability to bind to the apoptotic protease activating factor 1 (Apaf-1) once released from the mitochondria, creating a protein designated apoptosome which will proceed to activate an initiator caspase, caspase-9, forming a complex which will trigger cell destruction.
The p53 pathway that culminates in apoptosis is triggered by ATM. Active p53 releases cytochrome c from the mitochondria [50], thus promoting binding to apaf-1 and subsequent activation of caspase-9, while in contrast protein kinase B (Akt) negatively regulates pro-apoptotic proteins, thus inhibiting the activity of cytochrome c/apaf-1/caspase-9 complex. (see model diagram 7.1).

Chapter 4

Signalling Transduction Pathways

To carry out biochemical signals throughout the cell, a sophisticated mechanism evolved which includes several pathways collaborating in the complex intracellular environment. In this section we explore two major signalling transduction pathways which are important to the cell cycle paradigm since they influence its regulation direct or indirectly.

Epidermal growth factors (EGF) and its family of receptor tyrosine kinases are important members of the mechanism that underlies signal transduction, gene expression regulation and cell proliferation induced by an extracellular growth factor. The members of the receptor tyrosine kinase receptor family contain an extracellular ligand-binding domain, a single transmembrane domain, and a cytoplasmic protein tyrosine kinase domain. The pathway of the signal from the extracellular source to the cell's nucleus, where the gene expression is preformed, is done in multiple layers forming the so called MAPK cascade signalling pathway.

I The MAPK cascade signalling pathway

MAPK Cascade signalling pathway (Mitogenic-Activating-Protein Kinase Cascade), is a main mechanism for protein synthesis motivated by extracellular signals. It depends on MAPKKK, MAPKK and MAPK whose phosphorylated form is the activated form.

Extracellular signals, also called Ligands, such as *Growth Factors*, bind to transmembrane receptors, whose cytosolic domain may be allosterically altered, enabling its phosphorylation, inducing the binding of Growth factor receptor-bound protein 2 (GRB2) molecule, activating it. Active GRB2 activates *Son of Sevenless* (SOS), which in turn phosphorylates Ras-GDP complex to Ras-GTP complex. The latter can then activate Raf (MAPKKK) by binding. Raf will proceed the mechanism by phosphorylating MEK (MAPKK). Activated MEK promotes the phosphorylation of ERK (MAPK). Finally, active ERK promotes the activation of transcription factors and subsequent migration to the nucleus where it will bind to DNA transcription sites, leading to protein synthesis [7].
Phosphorylated ERK promotes cell growth [22]. Important transcription factors are the

Early Response Genes (ERG) *c-Fos*, part of the Fos family of transcription factors, the protein *c-jun* and the protein *Myc*. This cascade would continue indefinitely while Ras-GTP complex continues active. This is why this cascade also induces the transcription of GAP (GTPase-Activating Proteins) regulatory proteins, which act like a switch off button, phosphorylating Ras-GTP complex back to Ras-GDP complex, inhibiting the rest of the cascade and thus stopping the synthesis of the specific proteins that the mitogenic signals triggered. It has been documented that the silencing of GAP proteins is related to some human cancers [24], since it leaves the regulation of Ras protein to chance, resulting in the deregulation of the concentration of Ras, and consequently of the whole Cascade [see figure 4.1].
Overall, this signalling pathway needs to be well regulated to avoid cancers, because if one of the biomolecules involved in it were to be mutated, it could imply the consistent transcription of proteins necessary for deregulated growth and division, and thus it is only natural the study of drugs that reverse the "on" or "off" states of these biomolecules for cancer treatment, such as in [27].

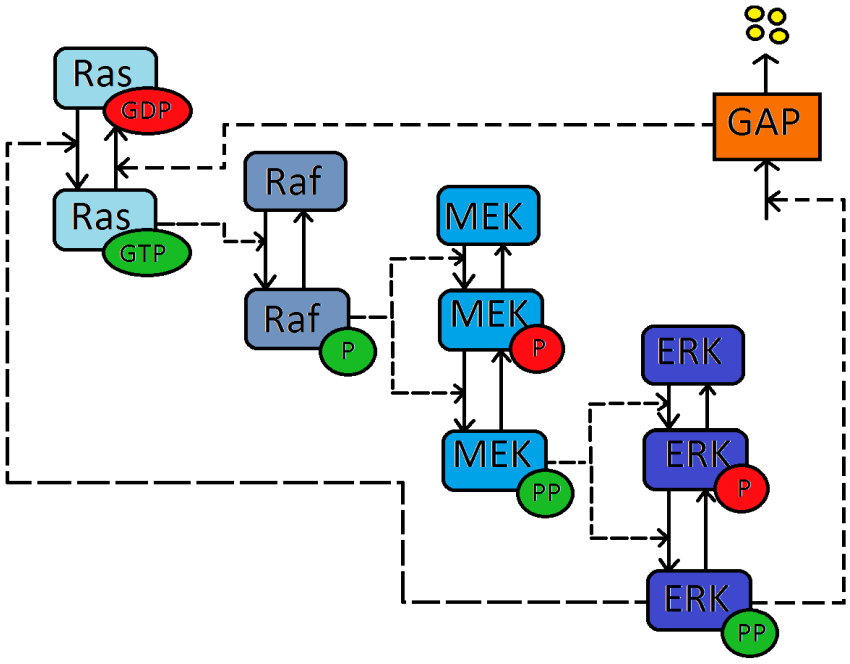

Figure 4.1: *The MAPK cascade signalling pathway*

II PI3K-AKT-mTOR pathway

Another important intracellular transduction pathway is the PI3K (phosphatidylinositol-4,5-bisphosphate 3-Kinase)-AKT-mTOR (mechanistic target of rapamycin) pathway (figure 4.2). This pathway is not as well studied as the MAPK cascade, however there are relevant dynamics that are sufficiently well documented, as the ones we describe in this section.
The regulatory subunit of PI3K, binds to phosphotyrosine peptide motifs in receptor protein

tyrosine kinases (RTK's) or the insulin receptor substrate 1 (IRS-1). This activates PI3K, which converts PIP2 (Phosphatidylinositol 4,5-bisphosphate) to PIP3 (Phosphatidylinositol (3,4,5)-trisphosphate), a reaction that is counter-acted by PTEN. PIP3 binds to AKT (Protein kinase B), forming the PIP3/AKT. PDK1 interveins to phosphorylate this complex, fully activating it. Active PIP3/AKT phosphorylates TSC2 (Tuberous sclerosis 2), deactivating it. TSC1 (Tuberous sclerosis 1) and active TSC2 form TSC1/TSC2 complex, which inhibit Rheb (Ras homolog enriched in brain) activity. In turn, Rheb promotes mTORC1 (mammalian target of rapamycin complex 1) activation. A feedback control in this PI3K/AKT pathway is the inhibition caused by mTORC1 in RTK dynamics, by phosphorylating IRS1, and also inhibiting EGF receptor, ERB2 and IGF1 receptor, therefore not allowing for interaction with PI3K and subsequent activation and downstream of the pathway [28].

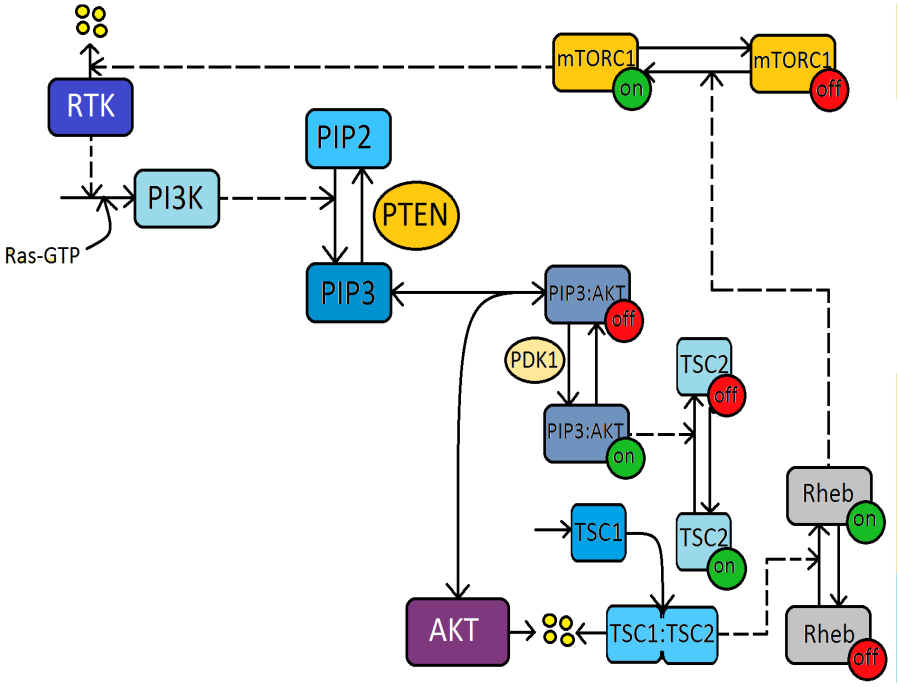

Figure 4.2: *The PI3K/AKT signalling pathway*

Chapter 5

Cancer: An overview

I Introduction

The biological revolution of the twentieth century, triggered by Watson and Crick's discovery of the DNA double helix, reshaped all fields of medical study, cancer research being one of them. Cancer started being studied as a genetic disease, which allowed researchers to link the genetic traits of cancer behaviour to intra- and inter-cellular pathways that come about. With the development of technology, these cellular pathways, were - and still are - being analyzed in evermore detail, promoting the establishment of new concepts in the field of biology, biochemistry and, specifically, cancer research.

Thanks to this improvement in scientific knowledge, cancer is now seen as a disease with a high level of complexity. The development of cancer is linked to many systems in complicated ways, and as such it comprehends a large number of degrees of freedom, making it particularly difficult to model due to the large number of variables that this implies. To say the same from a biological point of view, the eukariotic cell's is not an isolated system, and so interacts with its environment and responds to input stimuli in complicated ways. These output responses of a cell are based in many pathways created by the motion and/or action of particles, biomolecules or organelles, inside and outside the cell's membrane, and are orchestrated by the cell's genetic information. This is what allows the cell to produce proteins which function in some part of its metabolism. However, the genetic information, arranged in the DNA, does not operate in a deterministic way, but rather in a strong stochastic way. This complicates the study of genetic diseases, and hence this complicates the study of cancer. Not only the DNA of a particular cell is initially formed from another cell whose DNA may not be exactly the same due to mutations, but it also keeps on changing throughout the cell's life according to its environment. These changes in genetic information can originate a cancer cell. And a cancer cell generates offspring with potentially the same anomalies it has.

Even though the problem of modeling cancer seems too complicated to be done using uniquely deterministic approach and no type of data science whatsoever, our approach to cancer is done in a more humble way. Instead of tackling the disease from all the possible directions, studying the space and time dynamics of each component of all the possible pathways and deriving stochastic parameters to count for the randomness present

in nature, i.e., modelling each and every feature of cancer, we decided investigating only a subset of the characteristics, preparing the ground for posterior research to introduce more complexity to our model. This takes us to the next section.

II The hallmarks of cancer

In the year 2000, Hanahan and Weinberg defined six hallmarks of most, if not all, cancers. They proposed that acquiring the capability for autonomous growth signals, evasion of growth inhibitory signals, evasion of apoptotic cell death, unlimited replicative potential, angiogenesis (formation of new blood vessels), and invasion and metastasis are essential for carcinogenesis. More recently, two enabling characteristics were added, these being genome instability and tumor-promoting inflammation, that are crucial for acquiring the six hallmarks, and two emerging hallmarks, reprogramming energy metabolism (metabolic stress) and avoiding immune destruction, were highlighted [see figure 5.1]. Reprogramming energy metabolism and avoiding immune destruction are considered as emerging hallmarks because their relationship to the other mentioned hallmarks requires further research, even though it is clear they are relevant to carcinogenesis.
Healthy cells need external signals, such as growth factors, to divide, and they respond to inhibitory signals to maintain homeostasis (most cells of the body are not actively dividing). They have an autonomous counting device to define a finite number of cell doublings after which they become senescent. This cellular counting device is the shortening of chromosomal ends, called telomeres, that occurs during every cycle of DNA replication. Normal cells also maintain their location in the body, and generally do not migrate if that is not part of their function (for example, blood cells). These cells depend on blood vessels to supply oxygen and nutrients, but the vascular architecture remains more or less constant in an adult, meaning that there is no induction of angiogenesis. A healthy cell is eliminated by apoptosis, often in response to DNA damage.
Cancer cells on the other hand are independent of growth signals, and acquired mutations can shorten the pathway of these signals, leading to unregulated growth. Also, they have the capacity of ignoring growth inhibitory signals. There is evidence to support the theory of immune surveillance which states that the immune system can recognize and eliminate cancer cells. Cancer cells may be able to interfere or pass unnoticed by the immune response of the body so as to avoid immune destruction. This is one of the emerging hallmarks mentioned previously. Contrary to normal cells, cancer cells maintain the length of their telomeres, by reconstruction via the activity of the protein *telomerase*, which results in unlimited replicative potential. Virtually all tumors contain inflammatory immune cells. Inflammation is an immune response that can help that acquisition of the core hallmarks of cancer, which is why it is considered an enabling characteristic. This happens, for example, when inflammatory cells provide growth factors and enzymes that promote angiogenesis and invasion. Metastasis, i.e., the movement of cancer cells to other parts of the body, is a major cause of cancer deaths. Alterations of the genome may affect the activity or levels of enzymes involved in invasion or molecules involved in intercell or extracellular adhesion. Cancer cells also include angiogenesis, needed for tumor survival and expansion. Altering the balance between angiogenic inducers and inhibitors can activate the angiogenic switch,

and cancer cells promote this event. Acquiring the core hallmarks of cancer usually depends on genomic alterations, which can be caused by external sources and/or faulty DNA repair pathways, contributing to genomic instability. Thus, genome instability and mutation is an enabling characteristic of cancer. Evading apoptotic signals, which are crucial to induce apoptosis, i.e., programmed cell death, is a feature of every cancer cell. Finally, cancer cells have the capacity of reprogramming their own energy metabolism, considered to be an emerging hallmark. This is a useful characteristic for maintenance of energy given the uncontrolled cell division cancer cells go through. The demand in fuel and byosynthetic precursors is matched thanks to the intermediation of glycolysis carried out by cancer cells even in the presence of oxygen.
There two are major types of mutated genes that contribute to carcinogenesis, which are the oncogenes and the tumor suppressor genes. Generally, an oncogene is a gene mutated such that its protein product is produced in abundance, or has increased activity and therefore acts in a dominant manner - i.e., the mutation is only required in one allele to express an effect - to initiate tumor formation. Tumor supressor genes code for proteins that play a role in inhibiting both growth and tumor formation. Loss of growth inhibition occurs when mutations cause a loss of function of these genes. Tumor supressor genes are mainly recessive genes, which means that usually it suffices that one allele is kept intact to inhibit growth. However, recent evidence suggests there is a mechanism for particular tumor suppressor genes, called haploinsufficiency, whereby only one mutated allele can lead to the cancer phenotype.

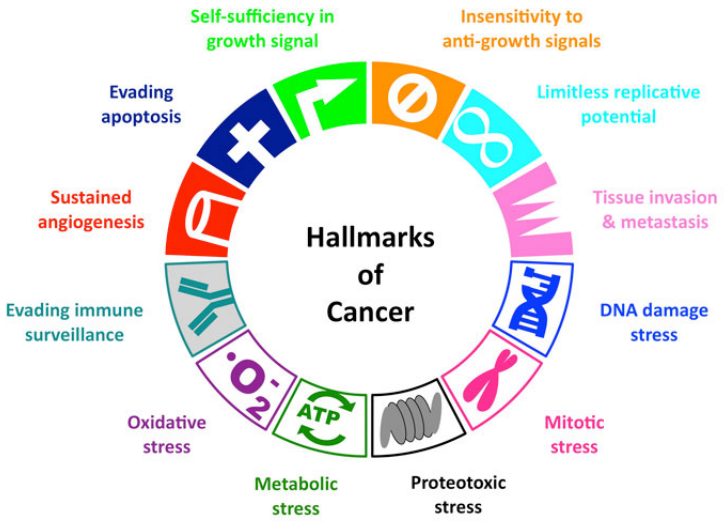

Figure 5.1: *The hallmarks of cancer*

III Relevant Pathways in Cancer and their deregulations

Even though cancer is a very complex disease, in the last decades, thanks to the development of technology, specific pathways and key elements of the cell have been identified and linked to carcinogenesis. The MAPK cascade signalling pathway and the PI3K-AKT-mTOR pathway are two major examples. We will explore how these two pathways, as well as key elements in the cell cycle, are related to cancer development.
The oncogenic activation of Ras is observed in about 30% of human tumors. This alteration of the Ras protein implies a consequent loss of GTPase activity of Ras. In normal circumstances, GTPase activity is required to return active Ras-GTP to inactive Ras-GDP. This leads to a constitutive activation of Ras protein, even in the absence of mitogens. Some specific mutations in the Ras gene are characteristic for specific cancers. The majority of mutations in Ras gene occur in codons 12, 13 and 61. A typical mutation within codon 12 that results in the substitution of valine (GTC) for glycine (GGC) is characteristic of bladder carcinoma, while substitution of serine (AGC) is common in lung cancer.
Another protein vulnerable to mutations is the B-Raf protein. The oncogenic form of B-Raf is common in melanomas. B-Raf's oncogenic activation causes kinase over-activity and insensivity to feedback mechanisms. MEK mutations, on the other hand, are less common: approximately 1% of tumors contain MEK mutations and no mutations have been identified as of 2010.
The PI3K pathway is commonly altered in colon cancer [51]. PI3K enhancement, AKT hyperactivity and PTEN loss are just some of the most common deregulations in this transduction pathway.
We already discussed in chapter 2 the Rb protein functions in the cell cycle. When this mechanism is corrupted, the cell can proceed to DNA replication even with DNA errors or other types of malfunctions, which can originate cancer. Rb is an indirect regulator of transcription for specific gene expression that affects cell proliferation and differentiation. Even though the Rb gene is expressed in all adult tissues, only retinoblastoma and a very few other types of cancer are initiated by loss of Rb. But still, the Rb pathway is inactivated in most human tumors.
A very important altered pathway in cancer is the p53 pathway, since it is deregulated in virtually all cancers. This high frequency may be the result of tumor cells that escape tumor suppressor effects of p53 through natural selection. p53 mutant cells are characterized by genomic instability. Over 75% of p53 mutations result in single amino acid substitutions, and in this sense p53 differs from other tumor suppressor genes, in that other tumor suppressors are usually characterized by nonsense or frame-shift mutations that lead to inactivated truncated proteins. Over-expression of the Hdm2 protein has been demonstrated to affect the regulation of p53, leading to a "p53-inactivated" phenotype.

IV Molecular targets in cancer therapies

The development in technology not only allows researchers to identify the origins of diseases but also to target core effectors in molecular pathways. This knowledge is vital for the field known as personalized medicine, which, in contrast to *one-fits-all* therapies,

investigates adequate individual treatments of diseases of each patient. As each cancer is intimately linked to genetic information of the host, it is not to wonder the progress that cancer therapies can have with the development of personalized medicine. These therapies target common deregulated pathways in human cancers, such as the ones discussed in the previous section.

Although 20% of all tumors have activating mutations in Ras, targeting downstream effectors has proved to be valuable as a cancer treatment. Several strategies to target Raf have been developed. The same idea used to target other kinases has been put into practice with Raf. A multi-kinase inhibitor that targets ATP-binding site of Raf, called NEXAVATAR^{TM} was approved for treatment of advance renal cell carcinoma and hepatocarcinoma in countries like the USA, Switzerland and Mexico. Results exhibited a reduction in downstream MAPK phosphorylation in the blood of patients under the well-tolerated oral treatment. Inhibiting MEK is a possible approach for Ras or Raf mutated patients. Several allosteric MEK inhibitors (inhibitors that do not compete with ATP) are in clinical trials.

The role of p53 as a core tumor suppressor and the high incidence of its mutations in cancer suggest promising p53-based therapies could be developed. The different ways p53 can be altered indicate that studying the p53 genotype prior to start a treatment could be the key for the success of p53-based therapies. Gene correction is therefore an obvious approach. Adenoviruses, such as *Onyx 015* adenovirus, have been used to selectively kill cancer cells with p53 mutations. Another way to target p53 pathway is to develop inhibitors of the p53-Hdm2 interaction since, as seen on chapter 3, Hdm2 inhibits p53. Diminishing tumor growth by 90% has been demonstrated in animal models following the idea of inhibiting p53-Hdm2 interaction. The success of chemotherapy and radiotherapy is often limited by side-effects in healthy tissue, and many of these side-effects are, in part, mediated by p53. There is normally high expression of p53 in tissues that are sensitive to these therapies and the DNA damage caused by them induces p53 to cause apoptosis, which is the mechanism behind the side-effects. Therefore, temporarily and reversible suppression of p53 in normal tissue may help alleviate the side-effects, but only in patients with tumors that have lost p53 function.

Chapter 6

State-of-the-art Models and Simulations

I Model and Simulation of Restriction Point regulation

Novak and Tyson constructed a deterministic mathematical model for the regulation of the restriction point based on the Cyclin/CDK complexes, the Rb-E2F interaction and the signal-transduction pathway, Growth Factors - Early response genes - Delayed response genes (GF-ERG-DRG) [3]. They assumed rapid message turnover for mRNA, that is, steady-state for mRNA transcription from the point of view of the other reactions, which occur in a longer timescale.
The GF-ERG-DRG is a brief pathway resuming the MAPK signalling Cascade that transducts extracellular signals through the cell's membrane and across the cytosol, reaching the nucleus and promoting the transcription of proteins. In this case, GF-ERG-DRG is going to be the triggering system that controls the CycD synthesis. CycD then binds to CDK4,6, also assumed to be fast enough to be in steady-state quickly, forming complexes CycD/CDK4,6 that phosphorylate Rb with the help of Cyclins A and B, inhibiting its action on E2F inhibition, therefore releasing free E2F. PP1 phosphatase promotes the dephosphorylation of Rb and is inhibited by CycE/CDK2, CycB/CDK1 and CycA/CDK2 complexes. CycE/CDK2 synthesis is in turn induced by E2F, and it can bind to Kip1, leading to the inhibition of CycE activity. This inhibiton by Kip1 is also applied to CycA. Kip1 degradation is mediated by CycB/CDK1 and CycA/CDK2. The former is mediated by Cdh1, that targets CycB for degradation. CycB, along with APC, promotes the deactivation of Cdh1, therefore establishing a mutual antagonism. Finally PPX mediates the synthesis of an intermediary enzyme, IE, whose phosphorylated form promotes the activation of CDC20. Active CDC20 and APC induce the degradation of CycA.
One of the most important features of this model is the growth and division simulation of the cell. The rate of mass is determined by the level of "General machinery", GM, controlled by the concentration of Rb and the absence of growth factors.
Using their model, they have simulated deregulations, such as cell lacking Rb, which

caused the loss of cell volume, and the continuing of the cell cycle beyond the restriction point. The model is an adaptation of their own yeast model, [4], to mammalian cells.

II Conradie Model and Simulation of Cell Cycle regulation

In the paper "*Restriction Point Control of the Mammalian Cell Cycle via the CycE/CDK2/p27 complex*", [8], the authors constructed a new framework in the restriction point model of Novak and Tyson, focusing on the CycE/CDK2/p27 complex. This model includes the dynamics of p27 and excludes the dynamics of Kip1 and Cyclins/Kip1 complexes. In this way, protein p27 becomes the cyclins' activity inhibitor – for every time p27 is active, the cyclins must be inactive. It does so by binding directly to the cyclins A,D,E complexes and keeping them inactive for the appropriate time in the cell cycle.
The changes in CycE/CDK2/p27 reactions lead to a shift in the restriction point, which allowed the authors to conclude that a perturbation in the concentration of this complex might cause the restriction point to happen earlier than expected.

III Kholodenko Model for the MAPK cascade pathway

The quantitative computational model done in [6] was performed around the MAPK cascade pathway, studying negative feedbacks, ultrasensitivity and emergent oscillations which simulate the nature of cellular biochemical pathways.
In this model, MKKKK (Ras) phosphorylates MKKK (Raf) to MKKK-P (Raf-P). In turn, MKKK-P phosphorylates MKK (MEK) to MKK-P (MEK-P) and MKK-P to MKK-PP (MEK-PP) which, finally, phosphorylates the bottom layer of the cascade, i.e., MAPK (ERK) to MAPK-P (ERK-P) and MAPK-P to MAPK-PP (ERK-PP). In this model, MKKKK (Ras) is not used as a variable, as its activity is simulated recurring to the negative feedback loop established by the downstream activity of ERK-PP. In this way, ERK-PP directly influences the concentration of MKKK (Raf).
It makes use of Michaelis-Menten kinetics [45] as main rate functions.

Chapter 7

Updated Mathematical Model of the Cell Cycle

The previous sections resumed the theory on which we based the construction of a wider updated mathematical [Figure 7.1], compared to that of Novak and Tyson.
In this new model, the MAPK signalling cascade was constructed as in [6], except for the fact that active ERK acts upon Ras-GDP phosphorylation, promoting it, instead of promoting Raf phosphorylation. This way the cascade has in its structure the dynamics of Ras, and maintains the feedback loop, even though Kholodenko used ERK's inuence on the activation of Raf [2]. Indirectly, this is still the case, as ERK is activating Ras that in turn activates Raf. The purpose of this was to propagate the oscillations created by the ERK feedback loop to Ras as well. Thanks to this, it was possible to add another negative feedback, through the inclusion of GAP protein, whose transcription is induced by the cascade and whose inhibition on Ras-GTP decreases the flux of the cascade. The activity of the cascade culminates in the regulation of ERG, whose influence on DRG models the transcription of CycD (see [3]).
The PI3K-AKT pathway was build inside another negative feedback loop - the inhibition of RTK by active mTORC1 against the activation of the latter by active Rheb. The joint concentration of active and inactive mTORC1 is not considered constant, as there is continuous synthesis and degradation of active mTORC1 (see equations (46) and (57)). Activity of AKT includes the formation of PIP3/AKT complex and the synthesis of CycD, which allows a simple connection to the cell cycle. There are several cross-talks between MAPK and PI3K-AKT pathways, and here we have focused only on two major influences acting upon PI3K-AKT pathway: activation of PI3K by active Ras and inhibition of TSC1/TSC2 complex by active ERK.
E2F promotes transcription of $p14^{ARF}$, suggesting another obvious link between models. As for p53, its concentration is maintained low through normal cell cycle by inhibition caused by Hdm2. The Hdm2 protein is downregulated by Rb, CycE/Cdk2 complex and p14, following [21]. The apoptosis and cell cycle arrest events were added with the features of the software used for the simulations, COPASI [48]. This is explained in more detail in

the next section (also, see notes on equations).
We based the core of our model, that is, the regulation of the Cyclins, in [8].

I Units of concentrations

The time dimension frequently used in intra-cellular activities is the $hour$, even though some reactions take longer than others. Taking $hour$ as the time unit, any reaction much faster than one hour (suppose a reaction of the order of seconds), can be seen as instantaneous, i.e., always in a steady state. For example, the time of messenger RNA (mRNA) turn over (passage from RNA to protein) is much faster than protein reactions time. From the time perspective of protein reaction, mRNA keeps roughly the same concentration throughout time, and therefore $\frac{d[mRNA]}{dt} \approx 0$, which does not really help much in the system of differential equations unless the constant $[mRNA]$ is used in a parameter somewhere in the equations. However, in this thesis the parameters are calibrated according to previous models and not directly from real-life dynamics.
To follow the modelling of Conradie (and also to be able to pick up the cell cycle model from Tyson and Novak) we assumed the concentration variables are scaled in order to have dimensionless Michaelis Menten parameters and rate constants with units $hour^{-1}$ [5]. The concentration units are in μM, where $M = mol/\mathrm{L}$.

II Simulation of Healthy Cell

As in [3], we implement in the model the cell cycle division, regulated by mass, whose concentration drops to half periodically (see notes on equations), as shown in the third plot of Figure 7.2 (left to right, top to bottom). The mass is the indicator of the current cell cycle phase. This is ought to depend on external factors, such as growth factors, but the whole simulation is going to be performed as if the healthy cell were in a stable environment, receiving periodical stimuli, and the cancer cell in a proper environment for its development. Therefore, we exclude external inputs for the model of both healthy and cancer cell, and can then focus on the internal regulation of cell. This takes us to the second link: ERG/DRG dynamics are an isolated system in Novak and Tyson model. As we discussed in the MAPK Cascade section, ERG activity is induced by transcription factors at the bottom of the MAPK Cascade, thus opening a hole in ERG/DRG isolation. For the sake of simplicity, we just assumed that the biphosphorylated form of ERK has a positive impact in ERG concentration. Inspired by Kholodenko [6], the model includes the MAPK cascade with a feedback created by the influence of active ERK on the phosphorylation of Ras-GDP. This establishes a loop around the MAPK cascade, generating one of the three sources of oscillation within this model - the others being the dynamics of CDC20 and CDH1 (second plot of Figure 7.2), and the negative feedback loop in PI3K-AKT pathway (fifth plot of Figure 7.2), which regulate the mass, which, in turn, regulates the cyclins (first plot of Figure 7.2). The Rb and E2F dynamics function as expected, following the cell cycle (sixth plot of Figure 7.2).
When it comes to the regulation of Hdm2, we assumed a constant flux of synthesis plus the

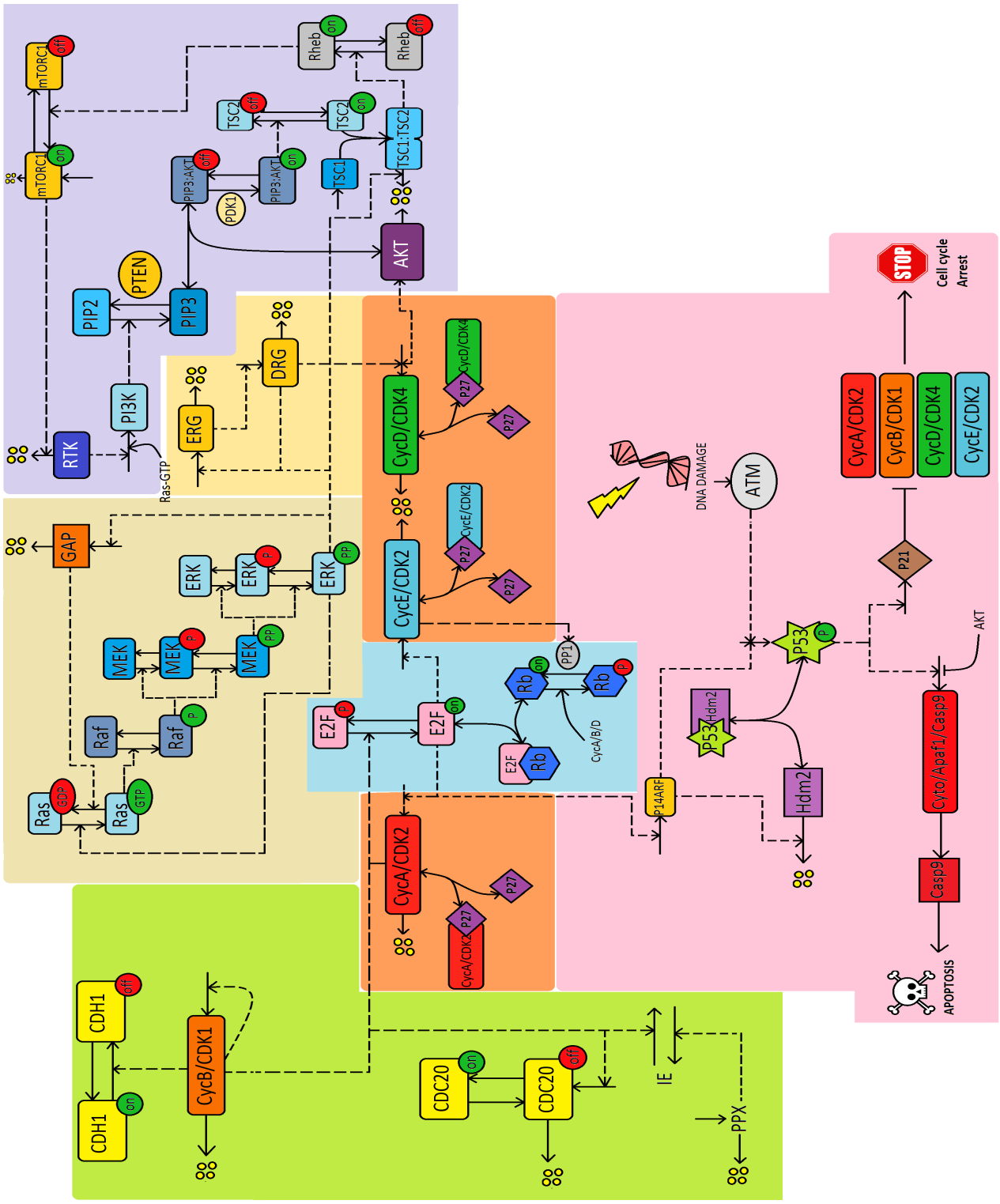

Figure 7.1: *Diagram of cell cycle dynamics divided according to specific functions in the model: generation of cell cycle oscillations through CDC20 and CDH1 reactions (green background); MAPK cascade (grey background), connection between MAPK cascade and production of Cyclin D through ERG and DRG synthesis (yellow background); PI3K/AKT pathway (light purple background); cyclin A/D/E and p27 interactions (orange background); restriction point regulation through E2F and Rb interactions (blue background); regulation of cell fate by p53 tumor supressor (pink background).*

induction by phosphorylated p53. The flux of synthesis is larger than the dependency on phosphorylated p53. Here the tumor suppressor is playing the role of regulating apoptosis and cell cycle arrest. Its concentration is low when Hdm2 is present, binding directly to it and increasing the concentration of Hdm2/p53 complex.
Simulating Cell Cycle Arrest implies a steady behaviour of cyclins concentration and cell volume for the time scale we are dealing with. It was therefore added a switch-like parameter in the rate functions of cyclins, General Machinery, CDH1 and mass, regulated by the condition of having a minimum amount of p21.
The programmed cell death can be triggered in two ways: from within the cell or through extracellular signals. Intracellular induction of apoptosis is the only one we concerned this project about, and therefore the intracellular pathway is the mechanism on which the simulation of apoptosis of the model is based on.
For the simulation of intracellular induction of apoptosis, it was added the conditions on minimum amounts of ATM which, if crossed, will trigger the increase in p53 dependency of Caspase-9. This induces a rapid construction of this protein, which will dismount the cell from within by degrading cyclins and leading the mass to zero in a switch-like way. This control of apoptosis and cell cycle arrest can be seen as very simple Boolean system.
All parameters in common to the mentioned models were maintained or only slightly changed. New necessary parameters were chosen according to the desired oscillation output and links between pathways, i.e., to allow the concentrations to sustain negative or positive feedback loops and at the same time establish smooth connections between distinct pathways of the cell cycle.
Notice that ERG, DRG and ERK concentration's order of magnitude is 100 times larger than the other ones. This is allowed to happen as long as the reaction parameters are adjusted to the correct value, in order to avoid over-dosing the system's reactions with these species. It was simpler to adjust the parameters to the correct flow (synthesis rate and degradation rate) of reactions than to alter initial concentrations, since Kholodenko's model was not obeying the same scale as Conradie's.
As a final note on the simulations, the number of hours in the x-axis are based on the assumption that a whole cycle of a healthy cell takes roughly 12 hours.

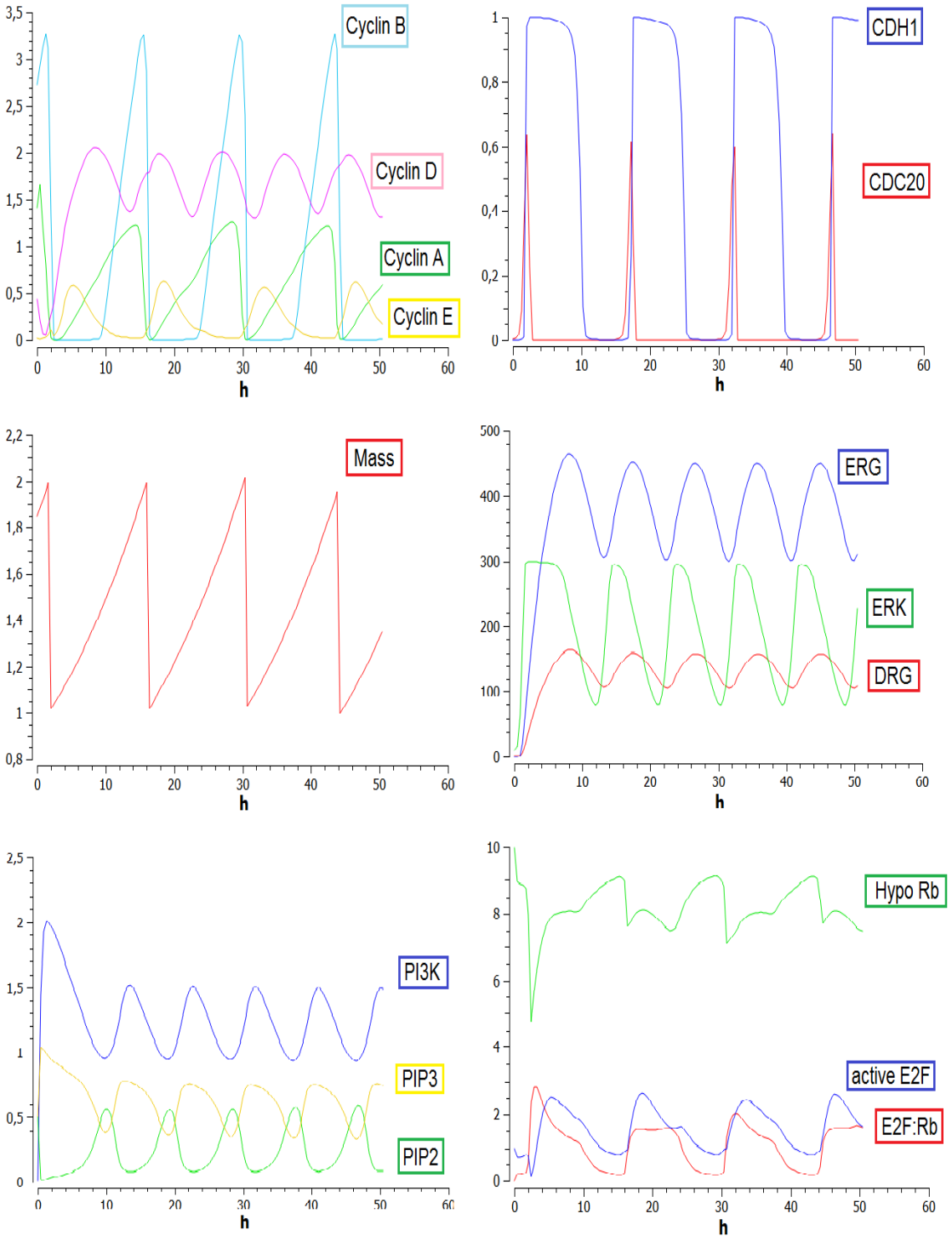

Figure 7.2: *Healthy cell simulations (details in text). Concentration unit: μM; Time unit: hour.*

Chapter 8

Simulation of Common Deregulations

After setting the model of a healthy cell, we can now proceed to simulate the cancer cell. Since cancer is a set of diseases rather than just one easily generalized disease, we will approach this simulation in two different ways: first, making use of the relevant information on MAPK cascade and PI3K/AKT pathway, on the p53 protein and its pathway, as well as the Retinoblastoma protein and the E2F transcription factor, to execute alterations on the model so as to simulate the beginning of a random cancer, which will deregulate some or several pathways. Second, starting in a specific type of cancer - colon cancer - we will alter the relevant pathways to then proceed to logical therapies described in literature.

I MAPK cascade signalling pathway in cancer

As a first approach to simulate cancer, the MAPK cascade dynamics in the model were altered, as it is a reasonable target for cancer study [7]. Starting by assuming mutated Ras or Raf proteins, such that its phosphorylating activity on Raf or MEK proteins, respectively, continues. For mutated Ras, the binding to GTP is still ongoing but the binding to GAP protein and inactivation by active ERK ceases to be possible - and thus Ras dephosphorylation form becomes rare in the cascade. For mutated Raf, the phosphorylation form becomes dominant and therefore the oscillation effect inherited by the Ras stage of the cascade is lost due to this overactivation of Raf. Both inefficient Ras-GTP hydrolysis and inefficient Raf dephosphorylation result in the same deregulation of the ERG production and therefore DRG as well [Figure 8.1]. The cyclins' periodicity overall does not suffer from this change, except that of CycD, whose amplitude diminishes, as we see in [Figure 8.2]. For severe dephosphorylation inhibition of both phosphorylated forms of ERK or severe dephosphorylation inhibition of biphosphorylated MEK, the effect on ERG, DRG and CycD are the same as with deregulation induced by Ras or Raf mutations [Figure 8.1]. In other words, as shown in figure 8.1, the same downstream result is obtained in one or more of the

above mentioned deregulations. The impact is not as strong for the analogous change on monophosphorylated MEK forms [Figure 8.3]. Thus, according to this simulation, it seems that ERK, compared to MEK, is a more sensitive component in the MAPK cascade, since any type of dephosphorylation inhibition on ERK suffices to loose the oscillation, while with MEK only if the biphosphorylated form suffers from severe phosphorylation inhibition.
A possible deregulation to explore in the MAPK cascade is the GAP synthesis, which inhibits directly active Ras. Increasing the GAP-dependency in the inactivation of Ras (which was done by increasing the parameter k_{gap} on equation (43)) and turning the production of GAP faster at the same time yields an increase in the number of waves of the MAPK cascade components' concentration [Figure 8.4] indicating a faster feedback. Since active ERK's concentration drops at the bottom of the cascade, there is a subsequent inhibition of the cell cycle, seen in Figure 8.5 (cyclins' concentrations are bellow 0.5 μM, which is quite low compared to healthy behaviour [see first graphic of 7.2]). p27 complexes are also affected by having its concentration frequency follow that of the MAPK cascade components (bottom graphic of Figure 8.6), in contrast to the healthy case (upper graphic of 8.6).

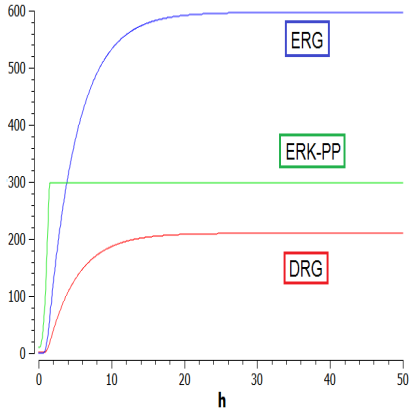

Figure 8.1: *Deregulation: inefficient Ras-GTP hydrolysis, Raf dephosphorylation or dephosphorylation of ERK-P*

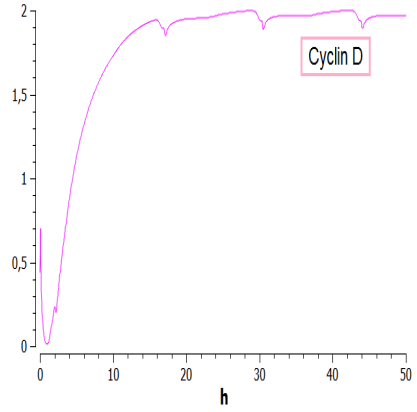

Figure 8.2: *Deregulation: inefficient Ras-GTP hydrolysis*

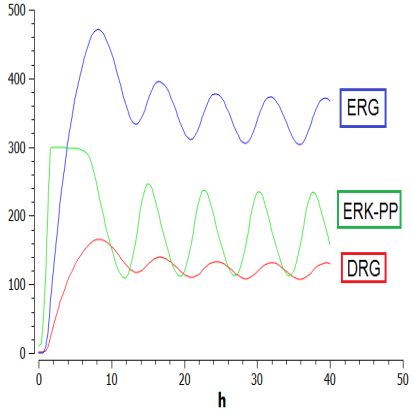

Figure 8.3: *Deregulation: inefficient dephosphorylation of MEK-P*

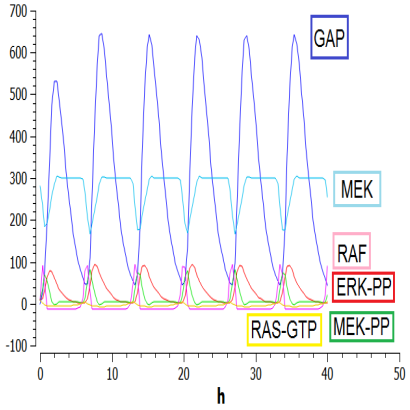

Figure 8.4: *Deregulation: strong dependency on and fast production of GAP*

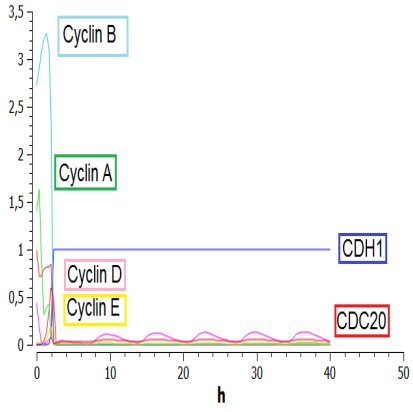

Figure 8.5: *Deregulation: strong dependency on and fast production of GAP*

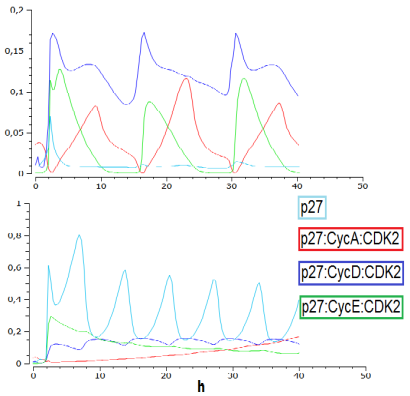

Figure 8.6: *Deregulation: strong dependency on and fast production of GAP*

II PI3K/AKT pathway deregulations

The PI3K pathway is commonly altered in colon cancer [40]. PI3K enhancement and PTEN loss are just some of the common deregulations in this transduction pathway. AKT hyperactivity is another possible deregulation.
For PI3K overexpression, synthesis was enhanced, leading to an ever-increasing concentration and loss of oscillation pattern in Rheb, mTORC1 and TSC1-2 concentrations [Figures 8.7-8.9]. Deregulations that are caused by the MAPK cascade, through cross-talks, are simulated in the colon cancer section below, due to their importance in this cancer.

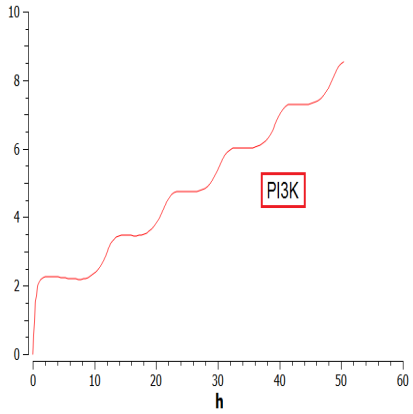

Figure 8.7: *PI3K overactivated*

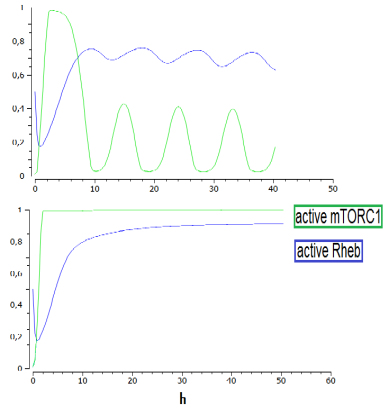

Figure 8.8: *Healthy case (upper plot) and deregulation: overactivated-PI3K (bottom plot)*

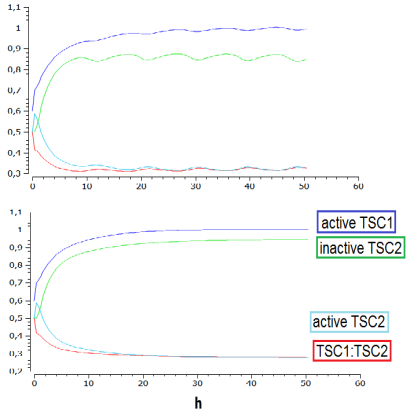

Figure 8.9: *Healthy case (upper plot) and deregulation: overactivated-PI3K (bottom plot)*

III CDH1 deregulation

Complete loss of CDH1 is implicated in 84% of lobular breast cancer [26], which may imply that the presence of CDH1 is a strong factor against lobular breast cancer. In the model, in absence of CDH1, i.e., with no synthesis whatsoever of CDH1, the wave length of the cycles in CDC20, the cyclins, p27 complexes, E2F and $p14^{ARF}$ are shortened (first four plots of Figure 8.10). As the mass only engages in division whenever CDH1 concentrations surpasses a certain threshold, the overall volume of the cell increases with no type of regulation (sixth plot of Figure 8.10). Although it may seem a positive feature of a deficient-CDH1 cell not having auto-induced division, like in cancer cells, this only happens due to the lack of cancer cell mechanism to induce its own division in this model. One can imagine that after CDH1 is removed completely from the cell, cell division is guided by another pathway, that is, a pathway which does not depend on CDH1 concentration nor CDC20 synthesis to induce cell division. Another outcome of this simulation is that p53 is not turned on (fifth plot of Figure 8.10), which means there is no apoptosis nor cell cycle arrest, and this combined with the already mentioned hidden cancerous mechanism for self-regulation of division becomes the perfect combination for the birth of a cancer cell.

IV Retinoblastoma mutation

If mutated, Rb can lose its ability to connect with E2F, not inhibiting its activity in proper time during the cell cycle, and consequently enabling E2F to migrate to the nucleus inducing the motion to S-phase in the cycle, moving past the restriction point. To simulate this, the chemical reaction of active Rb binding to E2F forming the complex E2F:Rb was shut off. This originated an increase in concentration of free E2F [Figure 8.11] and also an anticipation of the cell cycle, as the phase of the oscillation gets shifted [Figure 8.12], showing that the cell commits to an extra cell cycle earlier than in the healthy case, i.e., the restriction point was shifted. As a result of that, the general machinery does not have the proper time to build up, leading to an ever-decreasing mass in each cycle [Figure 8.13].
It is important to highlight that a similar result to the one observed in figure 8.12 comes about for alterations in the reactions $p27 + CycE : cdk2 \rightarrow p27 : CycE : cdk2$ and $p27 + CycE : cdk2 \leftarrow p27 : CycE : cdk2$, as reported in [8].

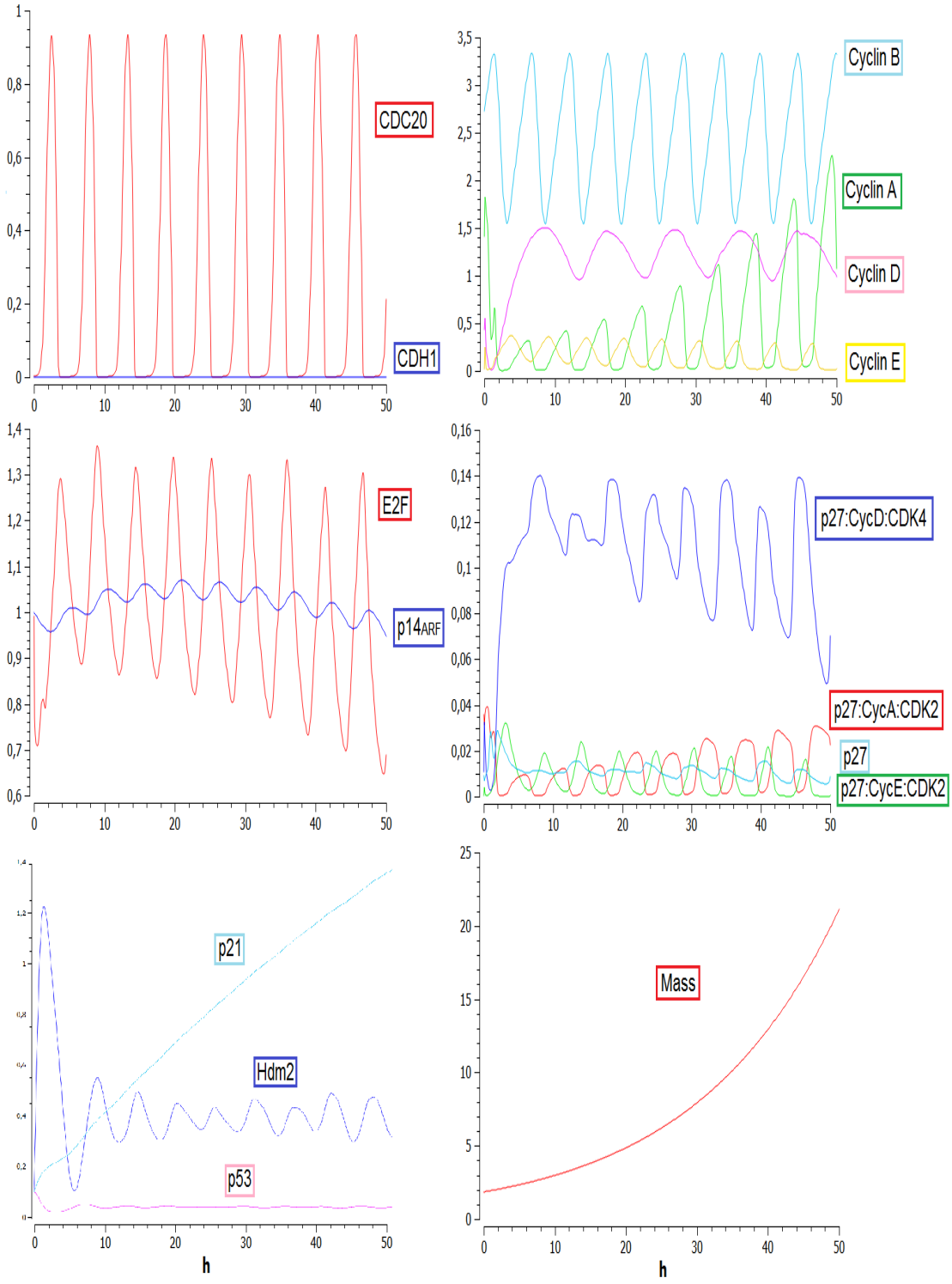

Figure 8.10: *Deregulation: CDH1 deficiency*

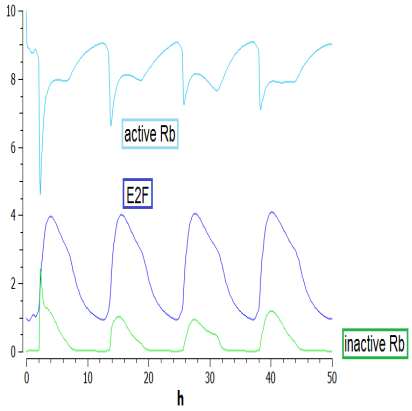

Figure 8.11: *Deregulation: Rb-mutated cell*

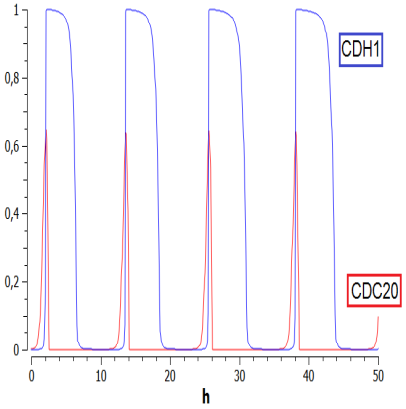

Figure 8.12: *Deregulation: Rb-mutated cell*

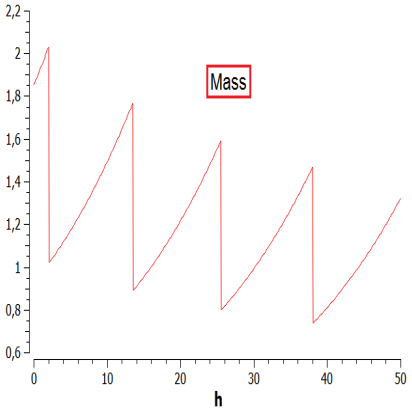

Figure 8.13: *Deregulation: Rb-mutated cell*

V P53 deregulated pathway

Any damage to the DNA is ought to induce the production of ATM, promoting p53 synthesis, which can trigger events for apoptosis or cell cycle arrest, the latter also involved in eventual repair mechanisms. If synthesis of p53 is deregulated, arising the appearance of mutated forms of p53, the ability of the cell to induce apoptosis or cell cycle arrest might be compromised. If mutated p53 is not able to promote p21 synthesis or Caspase-9 synthesis, then the cell does not have any defense mechanism, and will eventually continue through the cycle, passing errors in DNA to the daughter-cells, leading to irreparable cases of tumorigenesis.
To simulate this, the modifying role played by p53 in p21 and Caspase-9 synthesis was unintensified and the level of ATM was increased. The purpose is to simulate DNA damage in a mutated p53 environment. The result is clear [Figures 8.14-8.16]: the cell could not proceed to cell cycle arrest nor apoptosis and therefore the division continued. The concentrations of Hdm2 quickly reaches a constant level, since it has not enough supply of p53 to bind with.

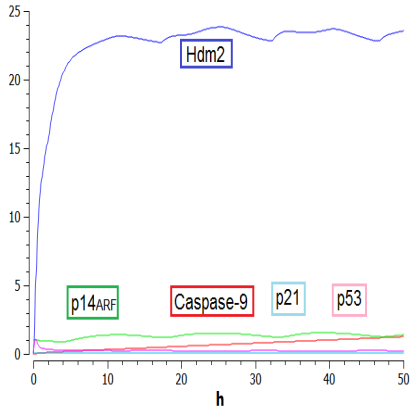

Figure 8.14: *Deregulation: mutated p53 such that it can no longer activate p21 synthesis*

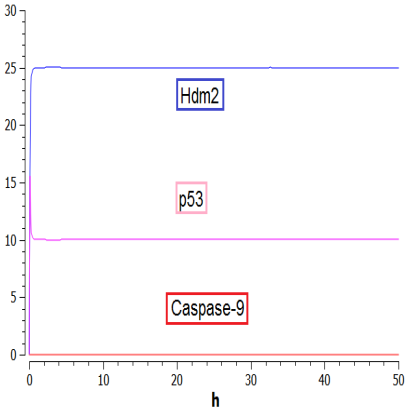

Figure 8.15: *Deregulation: mutated p53 such that it can no longer activate Caspase9 synthesis*

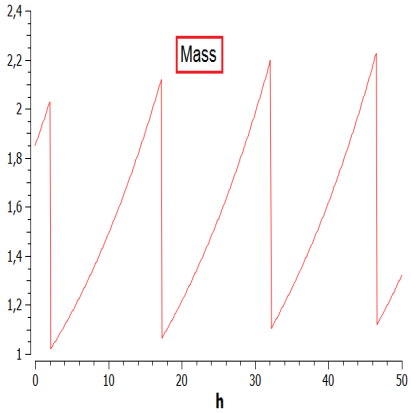

Figure 8.16: *Deregulation: p53-mutated cell*

Chapter 9

Colon Cancer Simulation

According to the World Cancer Research Fund International, colorectal cancer, or colon cancer, is the third most common type of cancer in the world and nearly 95% of colorectal cancers are adenocarcinomas, i.e., abnormal growth of epithelial tissue with glandular origin [51]. The molecular pathways our model simulates, are reasonable targets for therapy of this type of cancer. The Ras/Raf/MEK/ERK cascade is deregulated in approximately 30% of all cancers, Ras alone being mutated in 36% [42] and B-Raf (a specific type of Raf) found mutated in 10% to 15% of colorectal cancers [41]. The PI3K catalytic subunit alpha (PI3KCA) mutations are implicated in about 32% of colorectal cancers [42]. Hyperactivation of AKT has been linked to tumorigenic development, increasing cell survival, and was proved to be vital for colon cancer stem cells [43]. The regulatory system of the cell cycle is also afflicted by Ras mutations, which come with raised ERK activity [42]. On the other hand, Raf inhibitors have shown to be promising in certain cancers, with clinically manageable effects [44].
There is a correlation between colon cancer stages of progression and specific additional mutations: a normal cell with loss of APC can lead to a small benign polyp, which, with k-Ras activation, can conduce the cell to a Class II adenoma. From this stage, if the cell suffers loss of chromosome 18q gene, developing a Class III adenoma, and a subsequent loss of p53, it becomes a malignant carcinoma. From this last stage it can metastasize [1] [2].

I Simulation of colon cancer

Ras hyperactivity in colon cancer was seen as a straight-forward approach. As in the previous section on common deregulations, Ras dephosphorylation rate was diminished sufficiently to affect the rest of the MAPK cascade and PI3K/AKT pathway. There is a subsequent over-activation of PI3K, leading to a fast increase and posterior stagnation of PI3K/AKT pathway active components in a high concentration level. DRG and ERG concentrations frequency increase notoriously, following active ERK overactivation [Figure 9.1]. This contributes to shorter wave-lengths on CDH1 and CDC20 concentrations [Figure

9.2]. Hence the restriction point is overcome one more time than in healthy cell conditions. Uncontrolled division, yet, as mentioned previously, is not sufficient to create a cancer cell, according to the hallmarks of cancer [39], and thus, Ras hyperactivity, along with P53 mutation, was performed in this simulation, in this way clearly avoiding the Boolean switch of apoptosis or cell cycle arrest, generated by overproduced Caspase-9 or p21, respectively.
AKT hyperactivity affects mainly PI3K/AKT pathway, and doesn't seem to interfere with cyclins A,B and E, however it is still troublesome for it stimulates the overall production of CycD. TSC1 and TSC2 inherit the hyperactivity and raise, losing oscillation, while PIP3 lowers significantly and PIP3:AKT raises [Figure 9.3], both in active and inactive forms. PI3K, RTK and mTORC1 concentration's amplitude increases. Cell cycle is advanced, as with Ras activation [Figure 9.2].

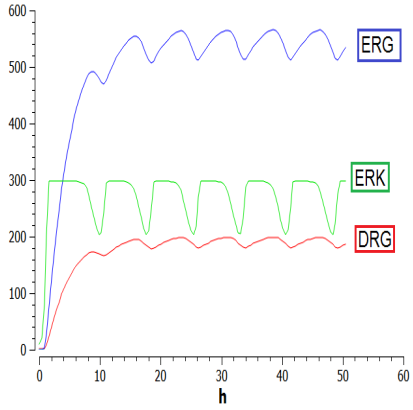

Figure 9.1: *Deregulation: hyperactive-Ras*

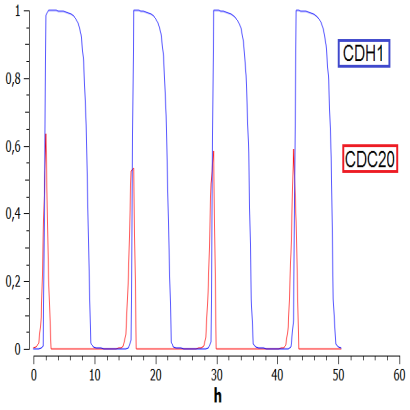

Figure 9.2: *Deregulation: hyperactive-Ras or hyperactive-AKT*

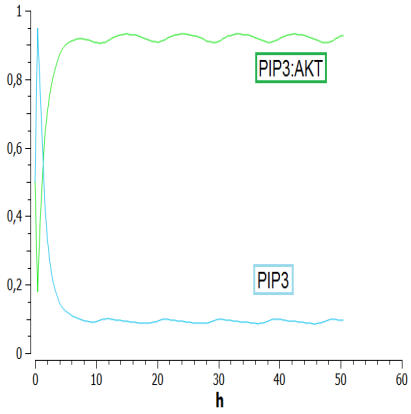

Figure 9.3: *Deregulation: hyperactive-AKT*

Chapter 10

Breast Cancer Simulation

According to the World Cancer Research Fund International (WCRF), breast cancer is the most common cancer afflicting women worldwide and the second most common cancer overall. In 2012, nearly 1.7 million new cases were diagnosed [52]. These facts, as well as the social impact of this disease, make this particular branch of cancer study a very important topic. Breast cancer is defined as an uncontrolled growth and division of breast tissue. The normal regulation of the cancer cells that make up that anomalous tissue has been corrupted at the level of their intra-molecular pathways.

I Simulation of breast cancer

Cell cycle main molecules have been linked to breast cancer in mutated forms or in abnormal concentrations, as we already discussed briefly in Cdh1 deregulation section. Cyclin D1 is overexpressed in many primary breast cancers [29] and it can lead, probably with the help of other oncogenes, to the development of mammary carcinoma [32]. This is because Cyclin D engages in other cancerous activities within the cell other than binding to Cdk4. This idea has been previously explored in 1994 [35]. Ras protein is often enhanced in breast cancer [29], [30], [31] which reveals a link to the MAPK cascade pathway. Cyclin D1 is a target of estrogen signaling [29]. Estrogen in turn is a hormone which binds to receptors in the cell membrane and promotes its growth and proliferation. About 70% of breast cancers, once established, rely on supplies of estrogen to grow, i.e., a tumor whose cells contain estrogen receptors (ER-positive), is fueled by estrogen supplies. This takes us to the study of the estrogen pathway [53]. Estrogen signaling pathway through the cell and into the nucleus relies on MAPK cascade as well as PI3K/AKT pathway. The deregulation of the PI3K/AKT and MAPK pathways has been hypothesized to sustain a metabolic switch, turning the former on and the latter off in a vast majority of breast cancers [36]. Inhibition of MAPK pathway shortens early G1, and activation of PI3K/AKT pathway induces the cell to progress to late G1, beyond the restriction point.
The simulation was performed based on the synthesized information from this section on breast cancer. It was added a estrogen-driven growth mechanism with the power of

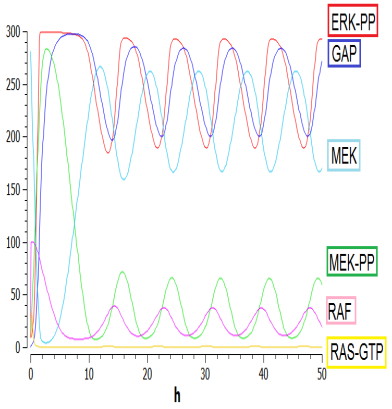

Figure 10.1: *MAPK cascade with estrogen-dependency*

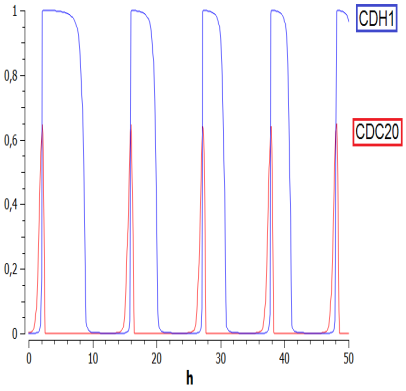

Figure 10.2: *CDH1 and CDC20 in estrogen-dependent cell model*

amplifying or reducing the cell volume rate. This was done by altering the synthesis rate of mass and that of Ras. Increasing the dependency on estrogen yields a faster volume increase [Figure 10.3] and shorter wave-lenghts in Cdh1, Cdc20 and MAPK cascade overal [Figures 10.1 and 10.2]. Assuming a stable cell growth, still estrogen-driven, the concentration rate of Cyclin D and Ras were deregulated. Corrupting the synthesis of Cyclin D by increasing its rate [Figure 10.6] as to create the impact of an unknown underlying oncogene pathway, the result is also an increase in Cdh1 and Cdc20 wave frequency [Figure 10.4], an unfair but expectable amount of CyclinD:p27 complexes [Figure 10.5], and an ever-decreasing cell volume [Figure 10.7]. No significant differences were seen in the complete lack of p27 protein, besides the obvious vanishment of p27 complexes. Cyclin regulation continued, only ever-so-slightly "disamplified". Even in the presence of ATM, conducing to larger amounts of p21, complete lack of protein p27 didn't seem to change the cell cycle significantly.

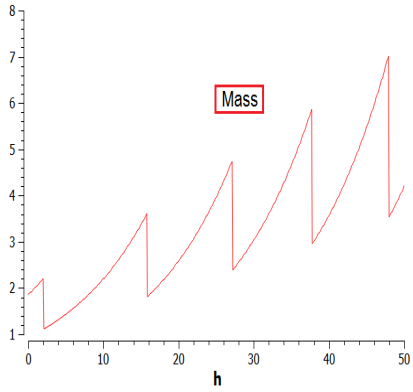

Figure 10.3: *Mass in estrogen-dependent cell model*

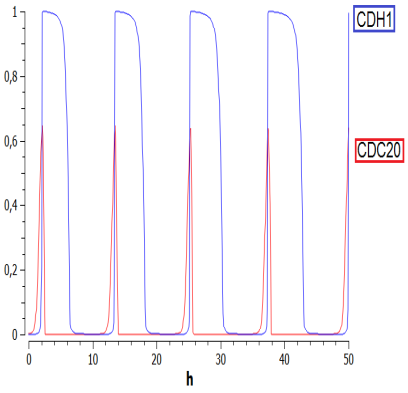

Figure 10.4: *CDH1 and CDC20 in overexpressed-Cyclin D cell*

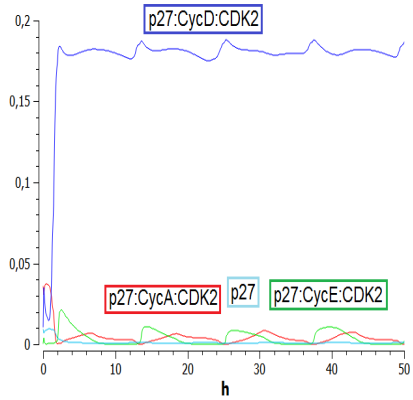

Figure 10.5: *p27 complexes in overexpressed-Cyclin D cell*

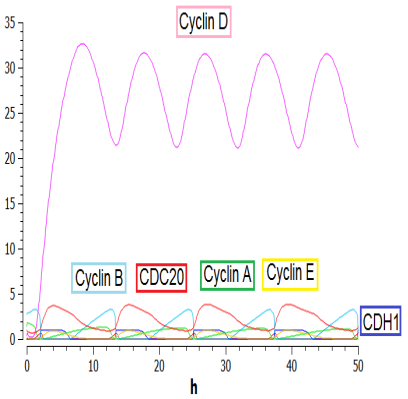

Figure 10.6: *Cyclins regulation in overexpressed-Cyclin D cell*

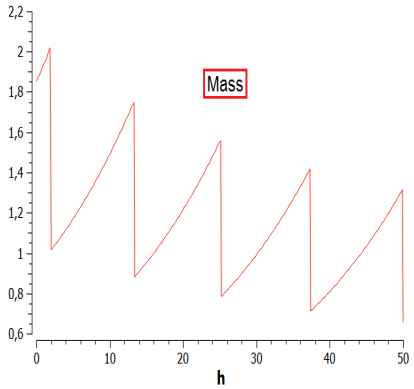

Figure 10.7: *Mass in overexpressed-Cyclin D cell*

Chapter 11

Therapies

The simulations performed wouldn't fulfill their purpose if one couldn't derive therapeutical intervention from them. Specifically, from the results it is possible to extract ideas for biomolecular-based therapies or, at least, to detect pre-cancerous cells.
MAPK cascade pathway's deregulation yields promising terrain for detecting possible cell cycle corruption: constant levels of active Ras, active Raf or active ERK can be measured indirectly through ERG and DRG concentrations [Figure 8.1] or Cyclin D concentration [figure 8.2]; as the increase in GAP-dependency leads to cyclin inhibition [Figure 8.5], targeting GAP in a previously-identified cancer cell for its overexpression is a possible way to inhibit cell growth.
Complete loss of CDH1 consequently induces the cell to a faster passage through the restriction point, whose regulation is done by E2F/Rb dynamics, meaning that a possible treatment for CDH1-deficient cells could be the artificial introduction of fair amounts of Rb or Rb-like biomolecules to create a delay that could ultimately compensate the fast pace of S-phase commitment; the same idea could be applied to Rb-mutated cells, since there are no biomolecules to hold the transcription factor activities of E2F in those cases.
Since p53 is mutated or its pathway is altered in virtually every cancer, therapies directly turned towards its activity in the cell could imply groundbreaking approaches to fight cancer. The pathway described in the model presented in this paper is not complex enough to allow one to easily derive therapies, but still some appear naturally: adding p53-like biomolecules to a p53-deficient cell, inducing p21 or Caspase-9 synthesis in cancerous cell or even targeting Hdm2 for destruction to indirectly increase the concentration of p53 are just some, rather simplistic, approaches. However, p53 pathway is not target-like, being extremely difficult to derive practical therapies from its deregulation.
It is also useful to make use of the altered model to find consequences of ideal therapies. With this in perspective, some simulation of therapies were performed, based in the common deregulation in colon cancer explored in previous sections. For Ras hyperactivation, MEK and Raf inhibition yield similar results, although the inhibitions were performed in different intensities: strong inhibition of Raf or inhibition of MEK induces cell cycle arrest without use of the p53 defense mechanism, starting with CDH1 and CDC20 stability [Figure 11.1]. The

cell signalling transduction ceases to operate, leading to stagnation of CycD production, and adding the arrest of CDH1 and CDC20, the cycle stops [Figure 11.2]. For this same deregulation, ERK inhibition treatment delays cell cycle [Figures 11.3 and 11.4], although not restoring completely MAPK cascade feedback [Figure 11.5] and leading to faster stationarity in PI3K/AKT dynamics. The delay seems to help increase the mass of the cell [Figure 11.6], since the first cell cycle takes longer to occur and the mass reaches higher values in the same period. As for PI3K inhibition, it does not seem very promising for Ras hyperactivity, because the cell cycle continues unharmed, even though PI3K/AKT components take longer to reach stationarity.
Even though these results seem promising, notice that a qualitative analysis would allow us to test the robustness of the model regarding therapy, by stretching the parameters in determined ranges to check how wide is the therapy effect. This because the model's parameters should be different for different patients, cell types and/or cancer types.

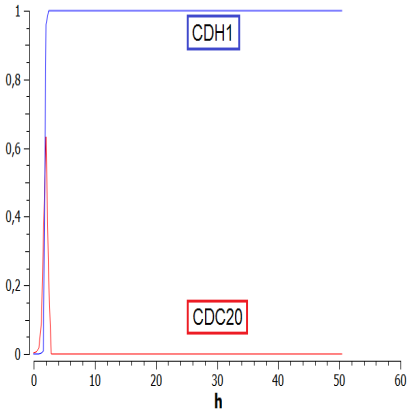

Figure 11.1: *Deregulation: hyperactive-Ras; Treatment: strong Raf inhibition or MEK inhibition.*

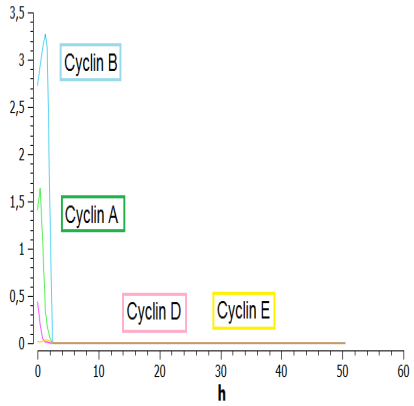

Figure 11.2: *Deregulation: hyperactive-Ras; Treatment: strong Raf inhibition or MEK inhibition.*

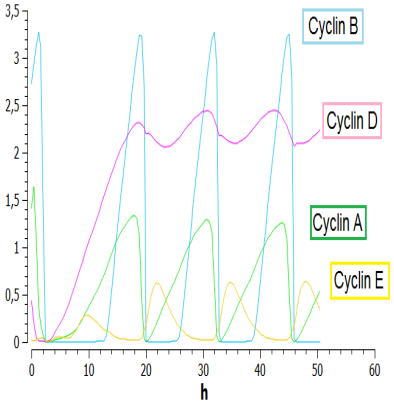

Figure 11.3: *Deregulation: hyperactive-Ras; Treatment: ERK inhibition.*

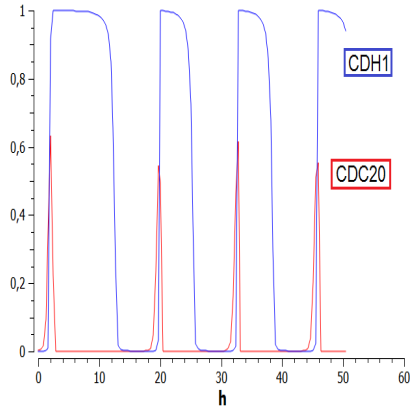

Figure 11.4: *Deregulation: hyperactive-Ras; Treatment: ERK inhibition.*

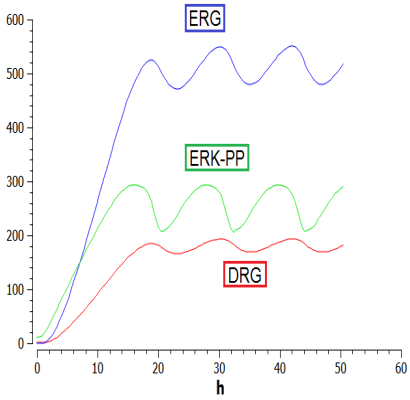

Figure 11.5: *Deregulation: hyperactive-Ras; Treatment: ERK inhibition.*

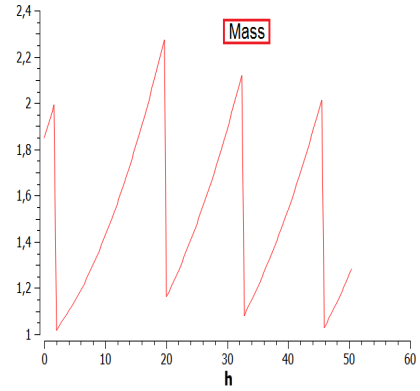

Figure 11.6: *Deregulation: hyperactive-Ras; Treatment: ERK inhibition.*

Conclusion

After the analysis of the contents of this thesis and the work in developing it, the conclusions drawn can be expressed in the following major topics:

First, the ease with which one can apply mathematics to real-life phenomena is indubitably one of the features that has made applied mathematics such a valuable field. Specifically, deriving systems of differential equations to model a network of chemical reactions occurring within a cell, has revealed to be simple enough under certain assumptions, in contrast with the difficulty of actually setting the reactions network in the first place. The main struggle in the development of this thesis was in fact the assembling of the knowledge in the fields of biology, biochemistry and oncology, to create a trustworthy model. These models' simulations run very fast in an ordinary PC, which is an advantage for any enthusiast in modelling. This is partially thanks to the existence of software like **COPASI** that allows to treat numeric and efficiently systems of differential equations. The other main reason why these simulations don't require much memory to run, is because the system is not very large, even though it can model fairly well important phenomena. For systems of partial differential equations, for which one can account space as another variable, the amount of information being processed is larger, resulting in time consuming simulations, despite the fact that these models are usually more accurate.

Second, creating extensions of already existing models can be a cumbersome experience, however it is a reliable method to create new models instead of building them from scratch. The model in the core of this thesis is still a simple model of the cell cycle, but it already begins dealing with cancer paradigms, and it was set in an easy-to-understand manner, which could help future researchers to expand it or enhance it according to more recent literature.

Third, many of the simulations results, reported in this thesis, mimic the most relevant and the most validated characteristics of statistical data extracted from many important *in vivo* or *in vitro* experiments, and clinical trials. This meaning that we did the effort to keep focus in the most corroborated results from literature. During the development of this thesis, it happen more often than not to be confronted with information that, even though promising, lacked support. In a few cases, information gathered from different sources was contradictory. However, these experiments are being updated everyday and faster than ever, which makes it particularly hard to be up-to-date.

The ideas for therapies explored in chapter 11 are only reasonable in a theoretical point of view, as time did not allow for further validation of the proposed treatments. The simulation of ideal target treatments for the specific case of Ras-hyperactivation already show that the model behaves as desired, inducing cell cycle arrest or regulating the cell.

In this thesis, we have began what can be later on the foundations of a system of partial differential equations model to simulate more accurately not only the cell but also a tissue of cells, bringing into the equations inter-cell interactions and eventually even simulate the core mechanism of metastasis. The simple equations of our model can easily be updated to match new information from *in vivo* or *in vitro* experiments.

Further work and investigation

The model presented in this thesis, even though it can already fit some of the experimental observations in literature, is extremely simplistic. For this reason it is important to mention approaches to construct a wider and more robust model of the cell and perhaps of cell-cell interaction.
To extrapolate the paradigm of the cell cycle used in this thesis to a wider model, more biomolecular species could be added, such as PCNA, CDC25, AP-1, c-fos, c-jun, myc and even ATP. This not only would require deeper understanding of how these species interact and influence the cell cycle, but also a careful examination of eventual groups of other biomolecules that couldn't be ignored by adding these new ones. It is also a risky move, as the complexity begins to increase fast with each new specie added.

Another idea is to explore more than one dimension. Instead of analyzing the concentration of molecules throughout time, space can also be part of the analysis. Space can be viewed with one, two or three dimensions, depending on the level accuracy we desire and the amount of work we are willing to do. The complexity of constructing such a model obviously increases with the number of spatial dimensions added. With one time dimension and one spatial dimension, we could already study the distance of each molecule to the nucleus of the cell, for example, while the cell cycle progresses. With one time dimension and two spatial dimensions it is possible to evaluate the proximity of each species to other organelles, such as the mitochondria, the Golgi apparatus and the endoplasmatic reticulum. Adding a third spatial dimension might be too much.
The units considered for concentration of intra-cellular molecules is usually $\frac{1}{12}$ of the Carbon-12 atom in ground state, $Dalton$ (Da), and not μM as it was used in this thesis and as previous authors used. Study appropriate concentration units (or checking if the ones used are the most reasonable) is also something to take into account in further work.
Detailed study of the most used target therapies should be a first step to begin deriving new target therapies to interact with the cell's biomolecular pathways. It is not only chemotherapy that could be a good field of research. Also radiotherapy and nanotechnology could be taken into account when modelling. Radiotherapy would be useful for a model that already accounts for cell-to-cell interactions. Nanotechnology is a field that is constructing very specific target therapies, and this is obviously an advantage to use in these mathematical models.
By constructing inter-cell models, it is already possible to simulate the activity of the body's immune system in attempts to stop tumor progression. A tumor can be simulated with a variety of heterogeneous cells (which in these types means each cell is deregulated with a

pathology (or several) from a wide spectrum of possibilities) and immunitary system's cells can be simulated indirectly by inhibiting some of the cell's tumors (the ones with the higher proportion of antigens in their membranes for example).
Cancer stem cells is another realm of possibilities in modelling that may not have been properly explored yet. In a tumor, these cells are less differentiated than their pairs and are the most dangerous, as they are the ones that are able to colonize and originate a new tumor in case of metastization.

System of differential equations for healthy cell model

$$\frac{d[Cdc20]}{dt} = \frac{k_{13}(-[Cdc20]+[''inactiveCdc20''])[''phosphorylatedIEP'']}{J_{13}-[Cdc20]+[''inactiveCdc20'']} - \frac{k_{14}[Cdc20]}{J_{14}+[Cdc20]} - k_{12}[Cdc20] \quad (1)$$

$$\frac{d[Cdh1]}{dt} = \frac{(k'_3 + k_3[Cdc20])(1-[Chd1])}{1+J_3-[Cdh1]} - \frac{V_4 CycleArrest[Cdh1]}{J_4+[Cdh1]} \quad (2)$$

$$\frac{d[CycA]}{dt} = k_{25r}[p27:CycA:Cdk2] + V_6[p27:CycA:Cdk2] + \epsilon k_{29} CycleArrest[mass][E2F] - k_{25}[CycA][p27] - k_{25}[CycA][Kip1] - k_{25r} - k_{30}[Cdc20][CycA] \quad (3)$$

$$\frac{d[CycB]}{dt} = \epsilon CycleArrest\left(k_{1'} + \frac{k_1[CycB]^2}{J_1{}^2(1+\frac{[CycB]^2}{J_{12}})}\right) - V_2[CycB] \quad (4)$$

$$\frac{d[CycD]}{dt} = V_6[p27:CycD:Cdk2] - k_{10}[CycD] + \epsilon CycleArrest k_9[DRG] + k_{akt}[AKT] - k_{24}[CycD][p27] + k_{24r}[p27:CycD:Cdk2] - k_{24}[CycD][Kip1] + k_{24r}[CycD:Kip1] + V_6[CycD:Kip1] + k_{10}[CycD:Kip1] \quad (5)$$

$$\frac{d[CycE]}{dt} = k_{25r}[p27:CycE:Cdk2] - V_8[CycE] + V_6[p27:CycE:Cdk2] + \epsilon(k_{7'} + k_7[E2F]) - k_{25}[CycE][p27] - k_{25}[CycE][Kip1] + k_{25r}[CycE:Kip1] + V_6[CycE:Kip1] + V_8[CycE:Kip1] \quad (6)$$

$$\frac{d[DRG]}{dt} = \epsilon\left(\frac{k_{17}[DRG]^2}{J_{17}{}^2(1+\frac{[DRG]^2}{J_{17}{}^2})} + k_{17}{}'[ERG]\right) - k_{18}[DRG] \quad (7)$$

$$\frac{d[E2F]}{dt} = k_{20}(\lambda_A[CycA] + \lambda_B[CycB] + \lambda_D([P27:CycD:Cdk2]+[CycD]) + \lambda_E[CycE])[E2F:Rb] + k_{26r}[E2F:Rb] - (k'_{23} + k_{23}([CycA]+[CycB]))[E2F] + k_{22}[''phosphorylatedE2F''] - k_{26}[E2F][Rb] \quad (8)$$

$$\frac{d[E2F:Rb]}{dt} = k_{26}[E2F][Rb] + k_{22}[''phosphorylatedE2F:Rb''] - k_{20}(\lambda_A[CycA] + \lambda_B[CycB] + \lambda_D([p27:CycD:Cdk2]+[CycD]) \quad (9)$$

$$
\begin{aligned}
& +\lambda_E[CycE])[E2F:Rb] - k_{26r}[E2F:Rb] - (k^{'}_{23} \\
& +k_{23}([CycA]+[CycB])[E2F:Rb] \\
\frac{d[ERG]}{dt} &= \frac{[ERG]k_{ERK}[''ERK-PP'']}{kk_{34}+[ERG]} + \frac{\epsilon k_{15}}{1+\frac{[DRG]^2}{J_{15}{}^2}} - k_{16}[ERG] \qquad (10)
\end{aligned}
$$

$$
\frac{d[GM]}{dt} = CycleArrestk_{27}[mass]r31switch - k_{28}[GM] \qquad (11)
$$

$$
\begin{aligned}
\frac{d[Rb\_hypo]}{dt} &= k_{20}(\lambda_A[CycA]+\lambda_B[CycB]+\lambda_D([p27:CycD:Cdk2]+[CycD]) \\
& +\lambda_E[CycE])([E2F:Rb]+[''phosphorylatedE2F:Rb'']+[Rb]) \\
& -(k_{19}PP1A+k_{19'}(PP1T-PP1A))[Rb_hypo] \qquad (12) \\
\frac{d[''inactiveCdc20'']}{dt} &= \epsilon(k_{11}{}^{'}+k_{11}[CycB]) - k_{12}[''inactiveCdc20''] \qquad (13) \\
\frac{d[mass]}{dt} &= \epsilon \cdot CycleArrest \cdot DeathSwitch\,\rho[GM] \qquad (14) \\
\frac{d[p27]}{dt} &= (k_{25r}+V_8)[p27:CycE:Cdk2] + (k_{25r}+k_{30}[Cdc20]) \\
& \cdot[p27:CycA:Cdk2] + (k_{10}+k_{24r})[p27:CycD:Cdk2] \\
& -(V_6+k_{25}([CycE]+[CycA])+k_{24r}[CycD])[p27]+\epsilon k_5 \qquad (15) \\
\frac{d[p27:CycA:Cdk2]}{dt} &= k_{25}[CycA][p27] - (k_{25r}+V_6+k_{30}[Cdc20])[p27:CycA:Cdk2] \qquad (16) \\
\frac{d[p27:CycD:Cdk2]}{dt} &= k_{24}[CycD][p27] - (V_6+k_{10}+k_{24r})[p27:CycD:Cdk2] \qquad (17) \\
\frac{d[p27:CycE:Cdk2]}{dt} &= k_{25}[CycE][p27] - (k_{25r}+V_6+V_8)[p27:CycE:Cdk2] \qquad (18) \\
\frac{d[''phosphorylatedE2F'']}{dt} &= (k_{20}(\lambda_A[CycA]+\lambda_B[CycB]+\lambda_D([p27:CycD:Cdk2]+[CycD]) \\
& +\lambda_E[CycE]+k_{26r}))[''phosphorylatedE2F:Rb''] \qquad (19) \\
& +k'_{23}+k_{23}([CycA]+[CycB])[E2F] \\
& -(k_{22}+k_{26}[Rb])[''phosphorylatedE2F''] \\
\frac{d[''phosphorylatedE2F:Rb'']}{dt} &= -k_{20}(\lambda_A[CycA]+\lambda_B[CycB]+\lambda_D([p27:CycD:Cdk2]+[CycD]) \\
& +\lambda_E[CycE]-(k_{26r}+k_{22}))[''phosphorylatedE2F:Rb''] \qquad (20) \\
& +k_{26}[''phosphorylatedE2F''][Rb] \\
& +(k'_{23}+k_{23}([CycA]+[CycB]))[E2F:Rb] \\
\frac{d[''phosphorylatedIEP'']}{dt} &= \frac{k_{31}[CycB](1-[''phosphorylatedIEP''])}{1+J_{31}-[''phosphorylatedIEP'']} \\
& -\frac{k_{32}[PPX]([''phosphorylatedIEP])}{J_{32}+[''phosphorylatedIEP'']} \qquad (21) \\
\frac{d[PPX]}{dt} &= \epsilon k_{33} - k_{34}[PPx] \qquad (22)
\end{aligned}
$$

$$\begin{aligned}\frac{d[Rb]}{dt} &= -(k_{20}(\lambda_A[CycA]+\lambda_B[CycB]+\lambda_D([p27:CycD:Cdk2]+[CycD])\\ &+\lambda_E[CycE]+k_{26r}))[Rb]+(k_{19}PP1A+k_{19}{}'(PP1T-PP1A))[Rb\_hypo]\\ &+k_{26r}[E2F:Rb]-k_{26}[E2F][Rb]+k_{26r}[''phosphorylatedE2F:Rb'']\\ &-k_{26}[''phosphorylatedE2F''][Rb] \qquad (23)\end{aligned}$$

$$\begin{aligned}\frac{d[Kip1]}{dt} &= \epsilon k_5 - V_6[Kip1]-k_{24}[CycD][Kip1]+k_{24r}[CycD:Kip1]\\ &+k_{10}[CycD:Kip1]+-k_{25}[Kip1]([CycE]+[CycA])\\ &k_{25r}([CycE:Kip1]+[CycA:Kip1])+V_8[CycE:Kip1]\\ &+k_{30}[Cdc20][CycA:Kip1]-((k_{25}([CycA]+[CycE])+k_{24}[CycD])[Kip1]\\ &-k_{25r}([CycA:Kip1]+[CycE:Kip1])-k_{24r}[CycD:Kip1]\\ &-V6([CycA:Kip1]+[CycD:Kip1]+[CycE:Kip1])\\ &-k_{30}[Cdc20][CycA:Kip1]-V8[CycE:Kip1]-k_{10}[CycD:Kip1]) \qquad (24)\end{aligned}$$

$$\begin{aligned}\frac{d[CycA:Kip1]}{dt} &= k_{25}[CycA][Kip1]-k_{25r}[CycA:Kip1]-V_6[CycA:Kip1]\\ &-k_{30}[Cdc20][CycA:Kip1] \qquad (25)\end{aligned}$$

$$\begin{aligned}\frac{d[CycE:Kip1]}{dt} &= k_{25}[CycE][Kip1]-k25r[CycE:Kip1]\\ &-V_6[CycE:Kip1]-V_8[CycE:Kip1] \qquad (26)\end{aligned}$$

$$\begin{aligned}\frac{d[CycD:Kip1]}{dt} &= k_{24}[CycD][Kip1]-k24r[CycD:Kip1]-\\ &V_6[CycD:Kip1]-k_{10}[CycD:Kip1] \qquad (27)\end{aligned}$$

$$\begin{aligned}\frac{d[Hdm2]}{dt} &= k_{37}[''p53-P''']+k_{38}-(k_{p14}[p14ARF]+k_{Rb}[Rb_hypo]\\ &+k_{cycE}[CycE])-\left(k_{39}[Hdm2][''p53-P''']-\frac{k_{40}[p53:Hdm2]}{kk_1+[p53:Hdm2]}\right)\end{aligned} \qquad (28)$$

$$\frac{d[p14ARF]}{dt} = k_{41}([E2F]-[p14ARF]) \qquad (29)$$

$$\frac{d[GAP]}{dt} = k_{ERK}[''ERK-PP''']-k_{42}[GAP] \qquad (30)$$

$$\frac{d[''p53-P''']}{dt} = K_{43}[ATM]-k_{39}[Hdm2][''p53-P''']+\frac{k_{40}[p53:Hdm2]}{kk_2+[p53:Hdm2]} \qquad (31)$$

$$\frac{d[p53:Hdm2]}{dt} = k_{39}[Hdm2][''p53-P''']+\frac{k_{40}[p53:Hdm2]}{kk_2+[p53:Hdm2]} \qquad (32)$$

$$\frac{d[p21]}{dt} = \frac{k_{44}[''p53-P''']}{k_{45}+[''p53-P''']}+\frac{k_{46}[PIP33:AKT\_On]}{kk_3+[PIP3:AKT\_On]}-k_{47}[p21] \qquad (33)$$

$$\frac{d[CytoC:Apaf1:Caspase9]}{dt} = ''Caspase9flux''[''p53-P'''] \qquad (34)$$

$$-k_{48}[AKT][CytoC:Apaf1:Caspase9] \tag{34}$$

$$\frac{d[Raf]}{dt} = \frac{k_{49}[''Raf-P'']}{kk_4+[''Raf-P'']} - \frac{k_{50}[''Ras-GTP''][Raf]}{kk_4+[Raf]} \tag{35}$$

$$\frac{d[''Raf-P'']}{dt} = \frac{k_{50}[''Ras-GTP''][Raf]}{kk_4+[Raf]} - \frac{k_{49}[''Ras-GTP''][Raf]}{kk_4+[Raf]} \tag{36}$$

$$\frac{d[MEK]}{dt} = \frac{k_{51}[''MEK-P'']}{kk_5+[''MEK-P'']} - \frac{k_{52}[''Raf-P''][MEK]}{kk_5+[MEK]} \tag{37}$$

$$\begin{aligned}\frac{d[''MEK-P'']}{dt} &= \frac{k_{52}[''Raf-P''][MEK]}{kk_5+[MEK]} - \frac{k_{52}[''Raf-P''][''MEK-P'']}{kk_5+[''MEK-P'']} \\ &+ \frac{k_{51}[''MEK-PP'']}{kk_5+[''MEK-PP'']} - \frac{k_6[''MEK-P'']}{kk_{15}+[''MEK-P'']}\end{aligned} \tag{38}$$

$$\frac{d[''MEK-PP'']}{dt} = \frac{k_{52}[''Raf-P''][''MEK-P'']}{kk_5+[''MEK-P'']} - \frac{k_{51}[''MEK-PP'']}{kk_5+[''MEK-PP'']} \tag{39}$$

$$\frac{d[ERK]}{dt} = \frac{k_{53}[''ERK-PP'']}{kk_5+[''ERK-P'']} - \frac{k_{52}[''MEK-PP''][ERK]}{kk_5+[ERK]} \tag{40}$$

$$\frac{d[''ERK-PP'']}{dt} = \frac{k_{52}[''MEK-PP''][''ERK-P'']}{kk_5+[''ERK-P'']} - \frac{k_{53}[''ERK-PP'']}{kk_5+[''ERK-PP'']} \tag{41}$$

$$\begin{aligned}\frac{d[''Ras-GDP'']}{dt} &= k_{gap}[GAP] + \frac{k_{54}[''Ras-GTP'']}{kk6+[''Ras-GTP'']} \\ &- \left(k_{mass}[mass] + \frac{k_{55}[''Ras-GDP'']}{\left(1+\left(\frac{[''ERK-PP'']}{ki}\right)^n\right)(k_{56}+[''Ras-GDP''])} \right)\end{aligned} \tag{42}$$

$$\begin{aligned}\frac{d[''Ras-GTP'']}{dt} &= k_{mass}[mass] + \frac{k_{55}[''Ras-GDP'']}{(1+(\frac{[''ERK-PP'']}{ki})^n)(k_{56}+[''Ras-GDP''])} \\ &-(k_{gap}[GAP] + \frac{k_{54}[''Ras-GTP'']}{kk6+[''Ras-GTP'']})\end{aligned} \tag{43}$$

$$\frac{d[RTK]}{dt} = k_{57} - k_{58}[RTK] - \frac{k_{59}[RTK][mTORC1]}{[RTK]+kk7} \tag{44}$$

$$\frac{d[PI3K]}{dt} = \frac{k_{Ras}[''Ras-GTP''][PI3K]}{kk8+[PI3K]} - \frac{k_{60}[PI3K]}{kk8+[PI3K]} \tag{45}$$

$$\begin{aligned}\frac{d[mTORC1]}{dt} &= \frac{k_{61}[mTORC1\_Off][Rheb\_On]}{kk8+[mTORC1\_Off]} + \frac{V3_{PI3K}(1-[mTORC1])}{kk8-[mTORC1]+1} \\ &- \frac{k_{62}[mTORC1]}{kk8+[mTORC1]} - \frac{k_{63}[mTORC1]}{kk9+[mTORC1]}\end{aligned} \tag{46}$$

$$\frac{d[PIP2]}{dt} = \frac{k_{63}[PI3K][PTEN]}{kk9+[PIP3]} - \frac{k_{63}[PI3K][PIP2]}{kk9+[PIP2]} \tag{47}$$

$$\begin{aligned}\frac{d[PIP3]}{dt} &= \frac{k_{63}[PI3K][PIP2]}{kk9+[PIP2]} \\ &+k_{64}[PI3K:AKT\_Off] - \frac{k_{63}[PIP3][PTEN]}{kk9+[PIP3]} - k_{65}[PIP3][AKT]\end{aligned} \tag{48}$$

$$\frac{d[PIP3:AKT\_Off]}{dt} = k_{65}[PIP3][AKT] - k_{64}[PI3K:AKT\_Off] - \frac{k_{63}[PDK1][PIP3:AKT\_Off]}{kk9+[PIP3:AKT\_Off]} + \frac{k_{63}[PIP3:AKT\_On]}{kk9+[PIP3:AKT\_On]} \quad (49)$$

$$\frac{d[PIP3:AKT\_On]}{dt} = \frac{k_{63}[PDK1][PIP3:AKT\_Off]}{kk9+[PIP3:AKT\_Off]} - \frac{k_{63}[PIP3:AKT\_On]}{kk9+[PIP3:AKT\_on]} \quad (50)$$

$$\frac{d[TSC2\_Off]}{dt} = \frac{k_{66}[''ERK-PP''][TSC2\_On]}{kk9+[''ERK-PP'']} - \frac{k_{64}[PIP3:AKT\_On][TSC2\_On]}{kk9+[TSC2\_On]} - \frac{k_{64}[TSC2\_Off]}{kk9+[TSC2\_Off]} \quad (51)$$

$$\frac{d[TSC2_On]}{dt} = \frac{k_{64}[TSC2\_Off]}{kk9+[TSC2\_Off]} + \frac{k_{67}[''ERK-PP''][TSC1:2]}{kk9+[''ERK-PP'']} - \frac{k_{66}[''ERK-PP''][TSC2\_On]}{kk9+[''ERK-PP'']} - \frac{k_{64}[PIP3:AKT\_On][TSC2\_On]}{kk9+[TSC2\_On]} - k_{67}[TSC1][TSC2\_On] \quad (52)$$

$$\frac{d[TSC1]}{dt} = k_{68} + \frac{k_{67}[''ERK-PP''][TSC1:2]}{kk9+[''ERK-PP'']} - k_{67}[TSC1][TSC2\_On] - k_{68}[TSC1] \quad (53)$$

$$\frac{d[TSC1:2]}{dt} = k_{67}[TSC1][TSC2\_On] - \frac{k_{67}[''ERK-PP''][TSC1:2]}{kk9+[''ERK-PP'']} \quad (54)$$

$$\frac{d[Rheb_Off]}{dt} = \frac{k_{69}[TSC1:2][Rheb\_On]}{kk9+[Rheb\_On]} - \frac{k_{70}[Rheb\_Off]}{kk9+[Rheb\_Off]} \quad (55)$$

$$\frac{d[Rheb\_On]}{dt} = \frac{k_{70}[Rheb\_Off]}{kk9+[Rheb\_Off]} - \frac{k_{69}[TSC1:2][Rheb\_On]}{kk9+[Rheb\_On]} \quad (56)$$

$$\frac{d[mTORC1\_Off]}{dt} = \frac{k_{64}[mTORC1]}{kk9+[mTORC1]} - k_{64}[mTORC1\_Off] - \frac{k_{68}[mTORC1\_Off][Rheb\_On]}{kk9+[mTORC1\_Off]} \quad (57)$$

$$\frac{d[AKT]}{dt} = k_{71} - k_{72}[PIP3][AKT] - k_{73}[AKT] + k_{64}[PIP3:AKT\_Off] \quad (58)$$

$$\frac{d[PDK1]}{dt} = k_{67}(1-[PDK1]) \quad (59)$$

$$\frac{d[PTEN]}{dt} = K_{74}(1-[PTEN]) \quad (60)$$

Steady-state equations

$$PP1A = \frac{[PP1T]}{1+K_{21}(\phi_E([CycE]+[CycA])+\phi_B[CycB])} \quad (61)$$

$$V_2 = k_2'(1-[CDH1]) + k_2[CDH1] + k_2''[Cdc20] \quad (62)$$

$$V_4 = k_4(\gamma_A[CycA] + \gamma_B[CycB]) \quad (63)$$

$$V_6 = k'_6 + k_6(\eta_E[CycE] + \eta_A[CycA] + \eta_B[CycB]) \quad (64)$$

$$V_8 = k'_8 \frac{k_8(\psi_E([CycE] + [CycA]) + \psi_B[CycB])}{J_8 + [CYCET]} \quad (65)$$

$$CYCET = [CycE : Kip1] + [CycE : Kip1] + [CycE] \quad (66)$$

$$CYCDT = [CycD : Kip1] + [p27 : CycD : Cdk2] + [CycD] \quad (67)$$

$$CYCAT = [p27 : CycA : Cdk2] + [CycA] + [CycA : Kip1] \quad (68)$$

$$P27T = [p27 : CycA : Cdk2] + [p27 : CycD : cdk2] + [p27 : CycE : Cdk2] + [p27] \quad (69)$$

$$V1_{PI3K} = \frac{[RTK]VM1_{PI3K}}{[RTK] + kc_{PI3K}} \quad (70)$$

$$V3_{PI3K} = [PI3K]VM3_{PI3K} \quad (71)$$

Notes on Equations:

(1) The mass of the cell drops to half, [mass] $\rightarrow$ [mass]/2, every time [Cdh1] crosses 0.2 from bellow.
(2) To simulate cell cycle arrest, whenever $[p_{21}] > 20$, $[CDH1] = [GM] = [CycA] = [CycB] = [CycD] = [CycE] = 0$. To do so, ***CycleArrest*** binary parameter was put in equations:

$$CycleArrest = \begin{cases} 0, & \text{if } [p21] > 20 \\ 1, & \text{in all other cases} \end{cases}$$

(3) As an event to prepare apoptosis, whenever $[ATM] > 20$, the synthesis rate of Caspase-9 increases with *"Caspase9 flux"* parameter:

$$"Caspase9flux" = \begin{cases} 5, & \text{if } [Cyto : Apaf1 : Caspase9] > 10 \text{ and } [ATM] > 20 \\ 10, & \text{if } [ATM] > 20 \end{cases}$$

(4) To simulate apoptosis, whenever $[Cyto : Apaf1 : Caspase9] > 10$ and $[ATM] > 20$, $[mass] = [GM] = 0$, using ***DeathSwitch***:

$$DeathSwitch = \begin{cases} 0, & \text{if } [Cyto : Apaf1 : Caspase9] > 10 \text{ and } [ATM] > 20 \\ 1, & \text{in all other cases} \end{cases}$$

(5) The same assumptions done in [3] were used in this system. The parameter ***r31switch*** works as a two steps Heaviside function:

$$r31switch = \begin{cases} 1, & \text{if } \frac{[Rb]+[E2F\_Rb]+["phosphorylatedE2F:Rb"]}{[Rb\_hypo]+[Rb]+[E2F:Rb]+["phosphorylatedE2F:Rb"]} < 0.8 \\ 0, & \text{in all other cases} \end{cases}$$

Parameters:

$n = 1$, $k_1 = 0.6$, $k'_1 = 0.1$, $k_3 = 140$, $k'_3 = 7.5$, $k_5 = 20$, $k_6 = 100$, $k'_6 = 10$, $k_7 = 0.6$, $k'_7 = 0$, $k_8 = 2$, $k'_8 = 0.1$, $k_9 = 0.05$, $k_{10} = 5$, $K_{10} = 3.8$, $k_{11} = 1.5$, $k'_{11} = 0$, $k_{12} = 1.5$, $k_{13} = 5$, $k_{14} = 2.5$, $k_{15} = 0.25$, $k_{16} = 0.25$, $k_{17} = 10$, $k'_{17} = 3.5$, $k_{18} = 10$, $k_{19} = 20$, $k'_{19} = 25$, $k_{20} = 10$, $k_{22} = 1$, $k_{23} = 1$, $k'_{23} = 0.005$, $k_{24} = 1000$, $k_{24r} = 10$, $k_{25} = 1000$, $k_{25r} = 10$, $k_{26} = 10000$, $k_{26r} = 200$, $k_{27} = 0.2$, $k_{29} = 0.05$, $k_{30} = 20$, $k_{31} = 0.7$, $k_{32} = 1.8$, $k_{33} = 0.05$, $k_{34} = 0.05$, $k_{35} = 5000$, $k_{36} = 10$, $k_{37} = 25$, $k_{38} = 2$, $k_{39} = 1$, $k_{40} = 0.5$, $k_{41} = 0.1$, $k_{42} = 1$, $k_{43} = 5$, $k_{44} = 100$, $k_{45} = 10$, $k_{46} = 10$, $k_{47} = 1$, $k_{48} = 0.1$, $k_{49} = 50$, $k_{50} = 100$, $k_{51} = 77.75$, $k_{52} = 2.855$, $k_{53} = 53$, $k_{54} = 1000$, $k_{55} = 1000$, $k_{56} = 10$, $k_{57} = 1$, $k_{58} = 0.5$, $k_{59} = 0.82$, $k_{60} = 0.15$, $k_{61} = 0.01$, $k_{62} = 1$, $k_{63} = 10$, $k_{64} = 0.5$, $k_{65} = 2$, $k_{66} = 0.7$, $k_{67} = 5$, $k_{68} = 0.1$, $k_{69} = 12$, $k_{70} = 4$, $k_{71} = 0.1$, $k_{72} = 2$, $k_{73} = 0.4$, $k_{74} = 1$, $k_i = 9$, $k_{AKT} = 0.5$, $k_{erk} = 0.5$, $k_{gap} = 0.1$, $k_{CycE} = 0.3$, $k_{p14} = 0.3$, $k_{Ras} = 0.6$, $k_{Rb} = 0.3$, $kk = 1$, $kk1 = 5$, $kk2 = 5$, $kk3 = 10$, $kk4 = 15$, $kk5 = 15$, $kk6 = 8$, $kk7 = 0.2$, $kk8 = 0.005$, $kk9 = 0.05$, $J_1 = 0.1$, $J_3 = 0.01$, $J_4 = 0.005$, $J_8 = 0.1$, $J_{13} = 0.005$, $J_{14} = 0.005$, $J_{15} = 0.1$, $J_{17} = 0.3$, $J_{31} = 0.01$, $J_{32} = 0.01$, $\epsilon = 1$, $\lambda_A = 3$, $\lambda_B = 5$, $\lambda_D = 3.3$, $\lambda_E = 5$, $\psi_E = 10$, $\psi_B = 0.5$, $\mu_E = 0.5$, $\mu_A = 0.5$, $\mu_B = 0.1$, $PP1T = 1$.

Glossary

14-3-3sigma is a human protein encoded by the SFN (stratifin) gene.

AKT is a protein kinase that plays a key role in glucose metabolism, apoptosis, cell proliferation, transcription and cell migration.

Allele is the short form of allelomorph (other form) used in genetics to describe variant forms of a gene detected with different phenotypes.

Allosteric regulation is the regulation of an enzyme by binding an effector molecule at a site other than the enzyme's active site. The site to which the effector binds is termed the allosteric site. Allosteric sites allow effectors to bind to the protein, often resulting in a conformational change involving protein dynamics.

Angiogenesis is the physiological process through which new blood vessels form from pre-existing vessel.

AP-1 (activator protein 1) is a transcription factor that regulates gene expression in response to a variety of stimuli, including growth factors, stress, and bacterial and viral infections.

Apaf-1 (Apoptotic protease activating factor 1) is a gene that encodes a cytoplasmic protein that forms one of the central hubs in the apoptosis regulatory network.

APC (Amino Acid-Polyamine-Organocation) is a family of transport proteins.

Apoptosis (from Ancient Greek "falling off") is a process of programmed cell death that occurs in multicellular organisms.

ATM (ataxia-telangiectasia mutated) is a serine/threonine protein kinase that is recruited and activated by DNA double-strand breaks. It phosphorylates several key proteins that initiate activation of the DNA damage checkpoint, leading to cell cycle arrest, DNA repair or apoptosis.

ATP (adenosine triphosphate) is a nucleoside and is a small molecule used in cells as a coenzyme. It is often referred to as the "molecular unit of currency" of intracellular energy transfer.

Carcinogenesis or oncogenesis or tumorigenesis is the formation of a cancer, whereby normal cells are transformed into cancer cells.

Caspase cysteine-aspartic proteases, cysteine aspartases or cysteine-dependent aspartate-directed proteases) are a family of protease enzymes playing essential roles in programmed cell death and inflammation.

Caspase-9 is an initiator caspase encoded by the CASP9 gene.

c-fos is a human proto-oncogene.

c-jun is a protein that in humans is encoded by the JUN gene. c-Jun in combination with c-Fos, forms the AP-1 early response transcription factor.

CDC2 (cell division cycle protein 2) is a highly conserved protein that has important functions in cell cycle regulation. It is also known as CDK1.

CDC20 (cell-division cycle protein 20) is an essential regulator of cell division that is encoded by the CDC20 gene in humans.

CDC25 (cell-division cycle protein 25) is a dual-specificity phosphatase.

CDH1 (Cadherin-1 or E-cadherin) is a protein that in humans is encoded by the CDH1 gene. It is a tumor suppressor gene.

Chromosome is a DNA molecule with part or all of the genome of an organism.

Cyclin is a biomolecule that controls the progression of cells through the cell cycle by activating cyclin-dependent kinase enzymes.

Cyclin A is a cyclin that activates CDK2.

Cyclin B is a cyclin that activates CDK1.

Cyclin D is a cyclin that activates CDK4 and CDK6. Growth factors stimulate the Ras/Raf/ERK that induce cyclin D production.

Cyclin E is a cyclin that activates CDK2.

CDK (cyclin-dependent kinases) are a family of protein kinases involved in regulating transcription and mRNA processing.

CDK1 (cyclin-dependent kinase 1)is a highly conserved protein that has important functions in cell cycle regulation. It is also known as CDC2.

CDK2 (cell division protein kinase 2) is an enzyme that in humans is encoded by the CDK2 gene.

CDK4/6 (cell division protein kinase 4 and cell division protein kinase 6, respectively), are two enzymes that in humans are encoded by the CDK4 and CDK6 gene, respectively.

Checkpoints are control mechanisms in eukaryotic cells which ensure proper division of the cell.

CKI (cyclin-dependent kinase inhibitor protein) is a protein which inhibits cyclin-dependent kinase.

Cytochrome c is a small human protein encoded by CYCS gene.

Cytosolic domain is the part of a transmembrane protein in interaction with the cytosol (cytoplasm) of the cell.

Cytotoxicity is the quality of being toxic to cells.

Coenzyme is a non-protein chemical compound or metallic ion that is required for a protein's biological activity to happen (for example, coenzymes such as ATP).

Degradation is the breakdown of proteins into smaller polypeptides or amino acids, also called proteolysis. It is typically catalysed by cellular enzymes.

DNA (deoxyribonucleic acid) is a molecule that carries the genetic instructions used in the growth, development, functioning and reproduction of all known living organisms and many viruses.

DRG (delayed-response genes) are genes which are activated at a slower rate than ERG.

E2F (E2 promoter binding factor) is a group of genes that codifies a family of transcription factors in higher eukaryotes.

Enzyme is a macromolecular biological catalyst.

ERG (early-response genes) are genes which are activated rapidly in response to a wide variety of cellular stimuli.

ERK is a protein kinase widely expressed in intracellular signalling pathways.

Eukariotic is any organism whose cells have a nucleus and other organelles enclosed within membranes.

Gadd45 (Growth Arrest and DNA Damage) is a protein implicated as stress sensors that modulate the response of mammalian cells to several types of stress, and modulate tumor formation.

GAP (GTPase-activating proteins) are a family of regulatory proteins whose members can bind to activated G proteins (proteins with the ability to bind GTP) and down regulate their activity

Genome is the genetic material of an organism.

Genotype is a DNA sequence which determines a specific characteristic (phenotype) of a cell/organism/individual.

GRB2 (Growth factor receptor-bound protein 2) is an adaptor protein involved in signal transduction/cell communication. In humans, the GRB2 protein is encoded by the GRB2 gene.

Growth factor is a naturally occurring substance capable of stimulating cellular growth proliferation, healing, and cellular differentiation. Usually it is a protein or a steroid hormone.

Hdm2 is a protein that in humans is encoded by the MDM2 gene. Since MDM2 means Mouse double minute 2 homolog, the human analog is sometimes refered to as Hdm2.

Kinase is an enzyme that catalyzes the transfer of phosphate groups from high-energy, phosphate-donating molecules to specific substrates.

Lymphocyte is one of the subtypes of white blood cell in a vertebrate's immune system.

MAPK (mitogen-activated protein kinase) is a type of protein kinase that is specific to the amino acids serine, threonine, and tyrosine.

MEK (mitogen-activated protein kinase kinase (MAPKK)) is a kinase enzyme which phosphorylates mitogen-activated protein kinase (MAPK).

Mole is the unit of measurement in the International System of Units (SI) for amount of substance. 1 mole is expressed by the Avogadro constant, which has a value of 6.022×10^{23}.

mTOR1 (mechanistic target of rapamycin 1) is a serine/threonine kinase mTOR which can inhibit the mechanistic target of rapamycin. It is a downstream effector of the PI3K/AKT pathway.

Myc (from "myelocytomatosis virus")is a regulator gene that codes for a transcription factor.

Oncogene is a gene related to the appearance of tumors. Usually, when not working properly, due to mutation for example, they can transform a healty cell into a cancer cell.

$p14^{ARF}$ (alternate reading frame protein 14) is a protein induced in response to elevated mitogenic stimulation. It acts as a tumor suppressor by inhibiting initiating p53-dependent cell cycle arrest or apoptosis.

$p16^{INK4a}$ is a tumor suppressor protein, that in humans is encoded by the CDKN2A gene.

p53 tumor protein, also known as TP53, is any isoform of a protein encoded by TP53 gene.

PCNA (proliferating cell nuclear antigen) is a DNA clamp essential for replication.

Phenotype is the composite of an organism's observable characteristics or traits.

Phosphorylation is the addition of a phosphoryl group $(PO_3)^-$ to a molecule.

PI3K (Phosphatidylinositol-4,5-bisphosphate 3-kinase) are a family of enzymes involved in cellular functions such as cell growth, proliferation, differentiation, motility, survival and intracellular trafficking.

PIP2 (Phosphatidylinositol 4,5-bisphosphate) is a minor phospholipid component of cell membranes.

PIP3 (Phosphatidylinositol (3,4,5)-trisphosphate) is a phospholipid that resides on the plasma membrane.

PP1 (protein phosphatase 1) belongs to a certain class of phosphatases known as protein serine/ threonine phosphatases.

Prokariotic is a unicellular organism that lacks a membrane-bound nucleus (karyon), mitochondria, or any other membrane-bound organelle.

Protein are large biomolecules, or macromolecules, consisting of one or more long chains of amino acid residues.

Proto-oncogene is a normal gene that could become an oncogene due to mutations or increased expression.

Raf (mitogen-activated protein kinase kinase kinase (MAPKKK)) is a kinase enzyme which phosphorylates mitogen-activated protein kinase kinase (MAPKK).

Ras (mitogen-activated protein kinase kinase kinase kinase (MAPKKKK)) is a kinase enzyme which phosphorylates mitogen-activated protein kinase kinase kinase (MAPKKK).

Rb (retinoblastoma protein) is a tumor suppressor protein that is dysfunctional in several major cancers.

Rheb (Ras homolog enriched in brain) is a GTP-binding protein that is ubiquitously expressed in humans and other mammals.

RNA (ribonucleic acid) is a polymeric molecule essential in various biological roles in coding, decoding, regulation, and expression of genes.

RTK (receptor tyrosine kinases) are the high-affinity cell surface receptors for many polypeptide growth factors, cytokines, and hormones.

SBML (Systems Biology Markup Language) is a representation format, based on XML (Extensible Markup Language), for communicating and storing computational models of biological processes.

SOS (son of sevenless) refers to a set of genes encoding guanine nucleotide exchange factors that act on the Ras family of small GTPases.

Synthesis is the execution of one or more named reactions to obtain a product, or several products.

Transcription factor is a protein that controls the rate of transcription of genetic information from DNA to messenger RNA, by binding to a specific DNA sequence.

TSC1 (tuberous sclerosis complex 1) is a protein that in humans is encoded by the TSC1 gene.

TSC2 (tuberous sclerosis complex 2) is a protein that in humans is encoded by the TSC2 gene.

umor suppressor gene are genes that code for fundamental proteins in the defense of a cell against tumor development. When this gene mutates causing a loss or reduction in its function, the

cell can progress to cancer, usually in combination with other genetic changes.

Ubiquitin is a small regulatory protein which can signal proteins for degradation, alter their cellular location, affect their activity, and promote or prevent protein interactions.

Ubiquitination is the addition of ubiquitin regulatory protein to a substrate protein.

Bibliography

[1] Robert A. Weinberg. The Biology of Cancer. Garland Science; 2 edition, 2013.

[2] Lauren Pecorino. Molecular Biology of Cancer: Mechanisms, Targets, and Therapeutics. OUP Oxford; 4 edition, 2016.

[3] Béla Novák, John J. Tyson, 2004. *A model for restriction point control of the mammalian cell cycle*. Journal of Theoretical Biology 230, 563-579.

[4] Chen. K.C., Csikask-Nagy, A., Gyorffy, B., Val, J., Novak, B., Tyson, J.J., 2000. *Kinetic analysis of a molecular model of the budding yeast cell cycle*. Mol. Biol. Cell 11, 369-391.

[5] Chen, Katherine C., et al. 2004. *Integrative analysis of cell cycle control in budding yeast.* Molecular biology of the cell 15.8: 3841-3862.

[6] Boris N. Kholodenko, 2000. *Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades*. Eur. J. Biochem. 267, 1583-1588.

[7] AS Dhillon, S Hagan and W Kolch. 2007 *MAP kinase signalling pathways in cancer*. Oncogene 26, 3279-3290.

[8] Riaan Conradie, Frank J. Bruggeman, Andrea Ciliberto, Attila Csikász-Nagy, Béla Novák, Hans V. Wsterhoff, Jacky L. Snoep, 2009. *Restriction point control of the mammalian cell cycle via the cyclin E/Cdk2:p27 complex*. FEBS Journal 277, 357-367.

[9] Charles J. Sherr, Frank McCormick, 2002. *The RB and p53 pathways in cancer*. Cancer Cell 2, 103-12

[10] Alberts B, Johnson A, Lewis J, et al. Molecular Biology of the Cell. 4th edition. New York: Garland Science; 2002.

[11] Zatterberg, A., Larsson, O., 1995. *Cell cycle progression and growth in mammalian cells: kinetic aspects of transition events*. In: Hutchison, C., Glover, D.M. (Eds.), Cell Cycle Control. Oxford University Press, Oxford, pp. 206-227.

[12] Peña, C., García, J.M., Silva, J., García, V., Rodríguez, R., Alonso, I., Millán, I., Salas, C., de Herreros, A.G., Muñoz, A., Bonilla, F. Hum. Mol. Genet. 2005, *E-cadherin and*

vitamin D receptor regulation by SNAIL and ZEB1 in colon cancer: clinicopathological correlations. Hum Mol Genet, 14(22):3361-70

[13] Kress M, May E, Cassingena R, May P. 1979, *Simian virus 40-transformed cells express new species of proteins precipitable by anti-simian virus 40 tumor serum*. Virol J 31: 472-483

[14] Lane DP, Crawford LV. 1979, *T antigen is bound to a host protein in SV40-transformed cells*. Nature 278: 261-263

[15] Linzer DI, Levine AJ. 1979, *Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells*. Cell 17: 43-52

[16] DeLeo AB, Jay G, Appella E, Dubois GC, Law LW, Old LJ. 1979 *Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse*. Proc Natl Acad Sci USA 76: 2420-2424

[17] Baker SJ et al. 1989 *Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas*. Science 244: 217-221

[18] Takahashi T, Nau MM, Chiba I, Birrer MJ, Rosenberg RK, Vinocour M, Levitt M, Pass H, Gazdar AF, Minna JD (1989) *p53: a frequent target for genetic abnormalities in lung cancer*. Science 246: 491-494

[19] Momand J., Zambetti GP, Olson DC, George D., Levine AJ. 1992, *The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation*. Cell 69(7):1237-45

[20] Picksley SM, Lane DP. 1993, *The p53-mdm2 autoregulatory feedback loop: a paradigm for the regulation of growth control by p53?* Bioessays 15(10):689-90

[21] Sandra L. Harris, Arnold J. Levine, 2005. *The p53 pathway: positive and negative feedback loops*. Oncogene 24: 2899-2908

[22] Meloche S., Pouysségur J. 2007, *The ERK1/2 mitogen-activated protein kinase pathway as a master regulator of the* G_1 *to S-phase transition*. Oncogene 26(22):3227-39

[23] Ameyar, M; Wisniewska, M; Weitzman, JB. 2003, *A role for AP-1 in apoptosis: the case for and against*.. Biochimie. 85 (8): 747-52.

[24] Jin, Hongchuan et al. *Epigenetic Silencing of a Ca2+-Regulated Ras GTPase-Activating Protein RASAL Defines a New Mechanism of Ras Activation in Human Cancers*. Proceedings of the National Academy of Sciences of the United States of America. 2007, 104: 12353-12358.

[25] Orton RJ, Sturm OE, Vyshemirsky V, Calder M, Gilbert DR, Kolch W (Dec 2005). *Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway*. The Biochemical Journal. 392 (Pt 2): 249-61

[26] De Leeuw WJ; et al. 1997. *Simultaneous loss of E-cadherin and catenins in invasive lobular breast cancer and lobular carcinoma in situ*. J Pathol. 183 (4): 404-411.

[27] Orton RJ, Sturm OE, Vyshemirsky V, Calder M, Gilbert DR, Kolch W (Dec 2005). *Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway*. The Biochemical Journal. 392 (Pt 2), 249-61

[28] Christoph Wagener, Carol Stocking, Oliver Müller, "*Cancer Signaling: From Molecular Biology to Targeted Therapy*". Wiley VCH, 2016.

[29] C. Elizabeth Caldon, Roger J. Daly, Robert L. Sutherland, and Elizabeth A. Musgrove, 2006, *Cell Cycle Control in Breast Cancer Cells*, Journal of Cellular Biochemistry 97:261-274

[30] Malaney S, Daly RJ. 2001. *The ras signaling pathway in mammary tumorigenesis and metastasis*. J Mammary Gland Biol Neoplasia 6:101-113.

[31] Schlessinger J. 2000. *Cell signaling by receptor tyrosine kinases*. Cell 103:211-225.

[32] Sutherland RL, Musgrove EA. 2004. *Cyclins and breast cancer*. J Mammary Gland Biol Neoplasia 9:95-104

[33] Musgrove EA, Davison EA, Ormandy CJ. 2004. *Role of the CDK inhibitor p27 (Kip1) in mammary development and carcinogenesis: Insights from knockout mice*. J Mammary Gland Biol Neoplasia 9:55-66.

[34] Bearss DJ, Lee RJ, Troyer DA, Pestell RG, Windle JJ. 2002. *Differential effects of p21(WAF1/CIP1) deficiency on MMTV-ras and MMTV-myc mammary tumor properties*. Cancer Res 62:2077-2084.

[35] Musgrove EA, Lee CS, Buckley MF, Sutherland RL. *Cyclin D1 induction in breast cancer cells shortens G1 and is sufficient for cells arrested in G1 to complete the cell cycle*. Proceedings of the National Academy of Sciences of the United States of America. 1994;91(17):8022-8026.

[36] Arnaud Guille, Max Chaffanet and Daniel Birnbaum, 2013, *Signaling pathway switch in breast cancer*, Cancer Cell International, 13:66

[37] Fresno Vara JA1, Casado E, de Castro J, Cejas P, Belda-Iniesta C, González-Barón M., 2004, *PI3K/Akt signalling pathway and cancer*. Apr;30(2):193-204

[38] Camillo Porta,1, Chiara Paglino, and Alessandra Mosca, 2014, *Targeting PI3K/Akt/mTOR Signaling in Cancer*, Front Oncol. 2014; 4: 64

[39] Hanahan D1, Weinberg RA., 2000, *The hallmarks of cancer*. Cell. 2000 Jan 7;100(1):57-70.

[40] Cathomas G. PIK3CA in Colorectal Cancer. Frontiers in Oncology. 2014;4:35. doi:10.3389/fonc.2014.00035.

[41] Corcoran RB, Ebi H, Turke AB, et al. EGFR-mediated re-activation of MAPK signaling contributes to insensitivity of BRAF mutant colorectal cancers to RAF inhibition with vemurafenib. Cancer discovery. 2012;2(3):227-235.

[42] Jing Yuan Fang, Bruce C Richardson, 2005. *The MAPK signalling pathways and colorectal cancer* . Lancet Oncol2005; 6: 322-27

[43] Y.K.Wang, Y.L.Zhu, F.M.Qiu, T.Zhang, Z.G.Chen, S.Zheng and J.Huang. *Activation of Akt and MAPK pathways enhances the tumorigenicity of CD1331 primary colon cancer cells.*Carcinogenesis vol.31 no.8 pp.1376-1380, 2010

[44] Libero Santarpia, Scott M Lippman and Adel K El-Naggar, 2012. *Targeting the MAPK-RAS-RAF signaling pathway in cancer therapy, Expert Opinion on Therapeutic Targets*, 16:1, 103-119

[45] L. Michaelis and Miss Maud L. Menten, 1913. *The Kinetics of Invertase Action*, translated by Roger S. Goody and Kenneth A. Johnson. Biochemistry, 2011, 50 (39), pp 8264-8269

[46] Rb: `https://www.ncbi.nlm.nih.gov/gene/5925`

[47] Conradie BioModels: `https://www.ebi.ac.uk/biomodels-main/BIOMD0000000265`

[48] Software used in simulations: `http://copasi.org/About/Team/`

[49] p53: `http://www.bioinformatics.org/p53/introduction.html`

[50] Schuler, Martin, et al. "p53 induces apoptosis by caspase activation through mitochondrial cytochrome c release." Journal of Biological Chemistry 275.10 (2000): 7337-7342.

[51] Colon cancer: `http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/colorectal-cancer-statistics`

[52] Breast cancer: `http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/breast-cancer-statistics`

[53] Estrogen pathway: `http://www.sabiosciences.com/pathway.php?sn=Estrogen_Pathway`

[54] Apoptosome, Cytochrome c and apaf-1: `https://en.wikipedia.org/wiki/Apoptosome`

[55] Hill equation generalization proof: `https://en.wikipedia.org/wiki/Hill_equation_(biochemistry)#cite_note-:0-6`